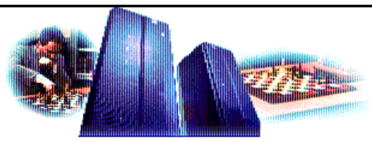
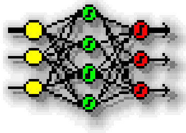




Introduzione al calcolo delle probabilità - II

Alberto Borghese

Università degli Studi di Milano
Laboratory of Applied Intelligent Systems (AIS-Lab)
Dipartimento di Informatica
borghese@di.unimi.it

Overview

Esempi sul teorema di Bayes

Densità di probabilità

Teorema di Bayes

$$P(X,Y) = P(Y|X)P(X) = P(X|Y)P(Y)$$

$$\boxed{P(X|Y) = \frac{P(Y \mid X)P(X)}{P(Y)}}$$

X = causa   Y = effetto

$$\boxed{P(\text{causa}|\text{effetto}) = \frac{P(Effetto \mid Causa)\, P(Causa)}{P(Effetto)}}$$

We usually do not know the statistics of the cause, but we can measure the effect and , through frequency, build the statistics of the effect or we know it in advance.

A doctor knows P(Symptons|Causa) and wants to determine P(Causa|Symptoms)

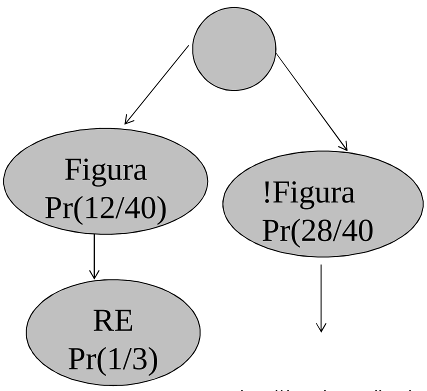
Probabilità condizionata

Consideriamo un mazzo di 40 carte:
vogliamo valutare quale sia la probabilità che una carta estratta a caso sia un re (probabilità semplice)
vogliamo valutare quale sia la probabilità che una carta estratta a caso sia un re, sapendo di avere estratto una figura (probabilità condizionata)

P(Y) = probabilità che sia un re
P(X) = probabilità che sia una figura

P(Y | X) = 1/3

P(Y) = P(Y/X) P(X) = 1/3 12/40 = 4/40



Esempio I

In una città lavorano due compagnie di taxi:
blue e verde: $X = \{T_{blu}, T_{verde}\}$ con una
Distribuzione di 85% di taxi verdi e 15% di taxi blu.

Succede un incidente in cui è coinvolto un taxi.

Un testimone dichiara che il taxi era blu. Era sera e l'affidabilità del testimone è stata valutata dell'80%.

Qual è la probabilità che il taxi fosse effettivamente blu?

Non è l'80%!

Esempio - I

$X = \{T_{blu}, T_{verde}\}$
"Causa"

$Y = \{C_{blu}, C_{verde}\}$
"Effetto"

Incidente

0.85 → Taxi verde
0.15 → Taxi blu

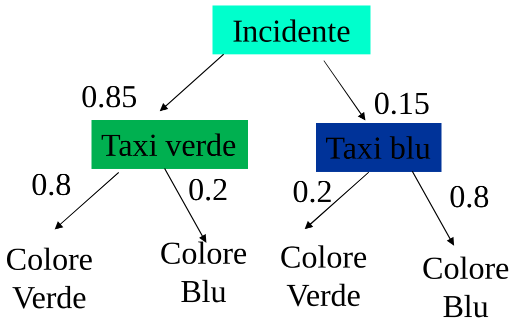
Taxi verde: 0.8 → Colore Verde; 0.2 → Colore Blu

Taxi blu: 0.2 → Colore Verde; 0.8 → Colore Blu

$$P(X=T_{blue}|Y_{blue}) = P(Y_{blue}|X_{blue})P(X_{blue}) / P(Y_{blue})$$

$P(X_{blue})$ = Probabilità a-priori = 0.15
$P(Y_{blue}|X_{blue})$ = Probabilità condizionata = 0.8

Esempio - I



$X = \{T_{blu}, T_{verde}\}$
$Y = \{C_{blu}, C_{verde}\}$

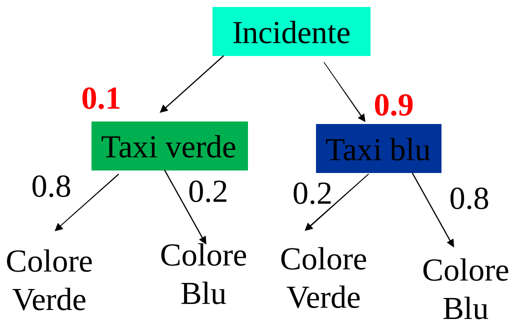
Inverto la relazione tra causa ed effetto applicando Bayes:
$P(X=T_{blu} \mid Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu})$

$P(Y_{blu})$ = Probabilità marginale di Y (probabilità semplice) =
$P(Y_{blu}|X_{blu})P(X_{blu}) + P(Y_{blu}|X_{verde})P(X_{verde}) = 0.8*0.15 + 0.2*0.85 = 0.29 > 0.15$!!

$P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.8*0.15 / 0.29 = 0.41$
Pesano anche gli “errori” commessi quando il testimone vede un taxi verde!

Esempio - I

Cambio la distribuzione dei taxi



$X = \{T_{blu}, T_{verde}\}$
$Y = \{C_{blu}, C_{verde}\}$

$P(X=T_{blu}|Y_{blu}) =$
$P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu})$

$P(Y_{blu})$ = Probabilità marginale di Y (probabilità semplice) =
$P(Y_{blu}|X_{blu})P(X_{blu}) + P(Y_{blu}|X_{verde})P(X_{verde}) = 0.8*0.9 + 0.2*0.1 = 0.74$

$P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.8*0.9 / 0.74 = 0.97$

Affidabilità della stima - SW in Matlab

N = 10 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.5$

N = 100 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.3235$

N = 1,000 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.4$

N = 10,000 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.4157$

N = 100,000 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.4173$

N = 1,000,000 ➔ $P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.4158$

Possiamo dare degli intervalli di confidenza?
Quanto deve essere grande N per ottenere una certa confidenza?

Estensione a più variabili

$P(X| Y_1;Y_2)$ if $(P(Y_1) = y_1$ and $P(Y_2) = y_2$ then $P(X) = x$

$Z = Y_1$ *and* Y_2

	mal di denti		!mal di denti	
	cavità	!cavità	cavità	!cavità
carie	0,108	0,012	0,072	0,008
!carie	0,016	0,064	0,144	0,576

Dalla tabella delle probabilità congiunte ricaviamo:

P(carie; Z) = P(mal di denti; cavità) = 0,108

P(carie | Z) = P(carie ; Z) / P(Z) = 0,108 / 0,124 = 0,871

Estensione a più variabili

$P(X| Y_1;Y_2)$ if $(P(Y_1) = y_1$ and $P(Y_2) = y_2$ then $P(X) = x$

$Z = Y_1$ *and* Y_2

	mal di denti		!mal di denti	
	cavità	!cavità	cavità	!cavità
carie	0,108	0,012	0,072	0,008
!carie	0,016	0,064	0,144	0,576

Applichiamo il teorema di Bayes

$P(X \mid Y_1$ and $Y_2) = P(Y_1$ and $Y_2 \mid X) * P(X)$

Abbiamo bisogno di conoscere come si comporta Z per ogni valore di X, cioè (Mal di denti *and* Cavità) in funzione di Carie. Diventa difficoltoso quando le variabili diventano tante: per capire se c'è una carie possiamo misurare anche: raggi-X, igiente orale....

Conditional independence

$P(X \mid Y_1$ and $Y_2) = P(Y_1$ and $Y_2 \mid X) * P(X)$

Introduciamo un'altra ipotesi. Cosa succede se Y_1 e Y_2 sono indipendenti? Dipendono entrambe da X ma non dipendono tra di loro.

Sono cioè **condizionatamente indipendenti**, cioè vale che:
$P(Y_1$ and $Y_2 \mid X) * P(X) = P(Y_1 \mid X) * P(Y_2 \mid X)$

In questo caso:

P(cavità and mal di denti | carie) = P(cavità | carie) * P(mal di denti | carie) che diventa più trattabile.

P(Carie ; cavità ; mal di denti) = P(Carie) [P(Cavità | Carie)*P(Mal di denti | carie)

Modello Naive Bayes Gli effetti sono indipendenti tra loro e dipendono da una stessa causa

In generale: $P(Causa \mid Effetto_1$ and $Effetto_2$ and ... $Effetto_N) = \boxed{\prod_{i=1}^{N} P(Effetto_i \mid Causa)}$

Conditional independence at work

$P(X| Y_1;Y_2)$ if $(P(Y_1) = y_1$ and $P(Y_2) = y_2$ then $P(X) = x$

	mal di denti		!mal di denti	
	cavità	!cavità	cavità	!cavità
carie	0,108	0,012	0,072	0,008
!carie	0,016	0,064	0,144	0,576

$P(X | Y_1$ and $Y_2) = P(Y_1$ and $Y_2 | X)$ * P (X) / $(P(Y_1)$ and $P(Y_2))$

P(carie | cavità and mal di denti) = P(cavità and mal di denti and carie) / P(cavità and mal di denti) = 0,108 / 0,124 = 87,1%

P(carie | cavità and mal di denti) = P(cavità and mal di denti | carie) * P(carie) ~~/ P(cavità and mal di denti)~~
P(carie | cavità and mal di denti) = P(cavità | carie) * P(mal di denti | carie) * P(carie) ~~/ (P(cavità) and P(mal di denti))~~

P(cavità | carie) = 0,18 / 0,2 = 0,9 P(carie) = 0,2
P(mal di denti | carie) = 0,12 / 0,2 = 0,6

P(carie | cavità and mal di denti) = (0,9* 0,6 * 0,2) / (0,124) = 87,1%

Esempio - II

Lo strumento principe per lo screaning per il tumore al seno è la radiografia (mammografia).

Definiamo X la situazione della donna: X={sana, malata}
Definiamo Y l'esito della mammografia: Y={positiva, negativa}

La sensitività della mammografia è intorno al 90%:

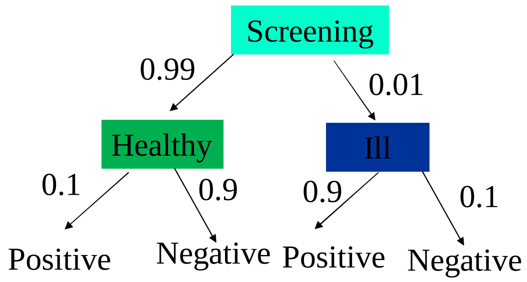
$$\text{sensitività} = \frac{n_{positive}}{N_{ill}} \quad \Rightarrow P(Y=positive \mid X=ill)$$

La specificità della mammografia è anch'essa intorno al 90%:

$$\text{specificità} = \frac{n_{negative}}{N_{healthy}} \quad \Rightarrow P(Y=negative \mid X=healthy)$$

Esempio II

X = {Healthy, Ill}
Y = {Positive, Negative}



P(Y=Positive | X=Ill) = 0.9 * 0.01 = 0.009
P(Y=Positive | X=Ill) = 0.1*0.99 = 0.099

P(Y=Positive) = P(Y=Positive | X=Ill) + P(Y=Positive | X=Healthy) =
0.009 + 0.099 = 0.108

10.8% di probabilità di avere un esame positivo a fronte di uno 0.01% di donne malate! Solo lo 0,9% proviene da donne effettivamente malate, le altre sono false positive

Esempio - II

Qual'è la probabilità che una donna sia veramente malata se il test risulta positivo?

Applichiamo Bayes:
P(X=Ill | Y=Positive) = P(Y=Positive | X=Ill)P(X=Ill) / P(Y=Positive)

P(X = Ill) = 0.01

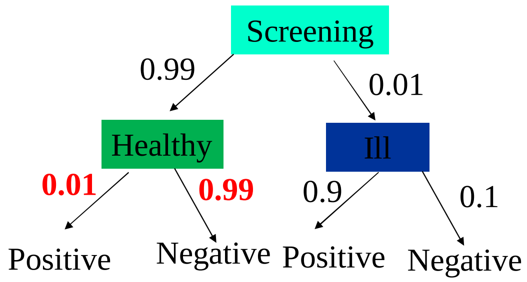
Il PPV (Positive Predictive Value) è:
P(X=Ill | Y=Positive) = P(Y=Positive | X=Ill)P(X=Ill) / P(Y=Positive) =
0.09 / 0.108 = 0.083 (8.3%)

Solo 8.3% delle donne con mammografia positiva sono effettivamente ammalate.

Analizzando la formula del teorema di Bayes, dove ha senso investire per ottenere un rendimento delle screening maggiore?

Esempio II

X = {Healthy, Ill}
Y = {Positive, Negative}



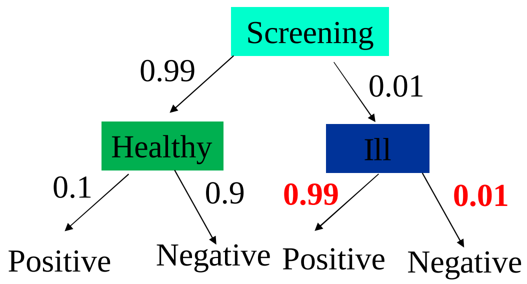
P(Y=Positive | X=Ill) = 0.9 * 0.01 = 0.009

P(Y=Positive = P(Y=Positive | X=Ill) + P(Y=Positive | X=Healthy) =
0.009 + 0.99*0.01 = 0.0189

P(X=Ill | Y=Positive) = P(Y=Positive | X=Ill)P(X=Ill) / P(Y=Positive) =
= 0.009 / 0.0189 = 0.476 = 47,6% >> 8.3%

Esempio II

X = {Healthy, Ill}
Y = {Positive, Negative}



P(Y=Positive | X=Ill) = 0.99 * 0.01 = 0.0099

P(Y=Positive = P(Y=Positive | X=Ill) + P(Y=Positive | X=Healthy) =
0.0099 + 0.99*0.1 = 0. 1098

P(X=Ill | Y=Positive) = P(Y=Positive | X=Ill)P(X=Ill) / P(Y=Positive) =
= 0.0099 / 0.1098 = 0.09 = 9% > 8.3%.

Riepilogo

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} = \frac{P(X,Y)}{P(Y)}$$

Teorema di Bayes

Lega probabilità condizionate, congiunte, semplici (marginali)

Consente di inferire la probabilità di un evento causa, X, a partire dalla probabilità associata alla frequenza di una certa misura, effetto, P(Y), dalla frequenza relativa dell'evento associato alla misura, P(Y), e dalla probabilità nota a-priori, P(X), della causa.

La probabilità P(X|Y) viene per questo detta probabilità a-posteriori ed è una probabilità condizionata.

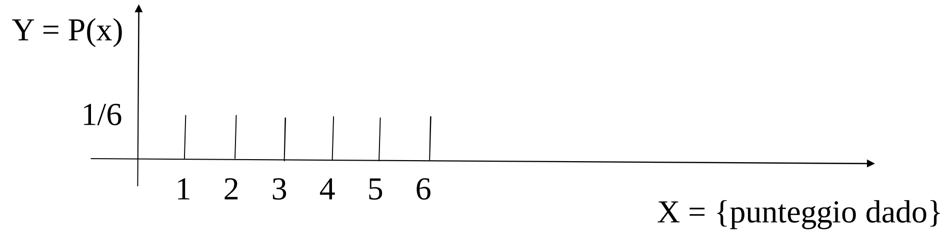
Viene utilizzata nei problemi inversi.

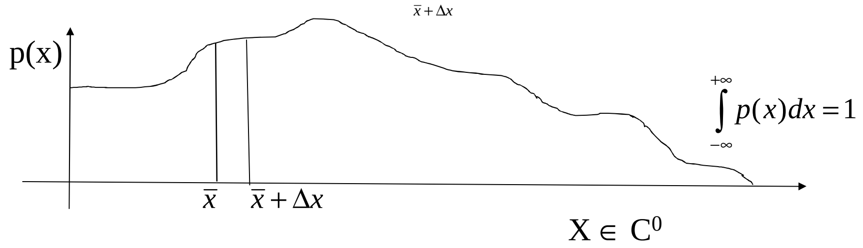
Overview

Esempi sul teorema di Bayes

Densità di probabilità

La probabilità nel caso continuo





$$P(x \in [\bar{x}, \bar{x} + \Delta x]) \int_{\bar{x}}^{\bar{x}+\Delta x} p(x)dx$$

Definizione di p(x)

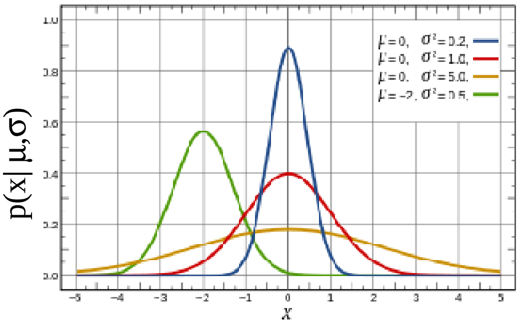
Caso discreto: prescrizione della probabilità per ognuno dei finiti valori che la variabile X può assumere: P(X).

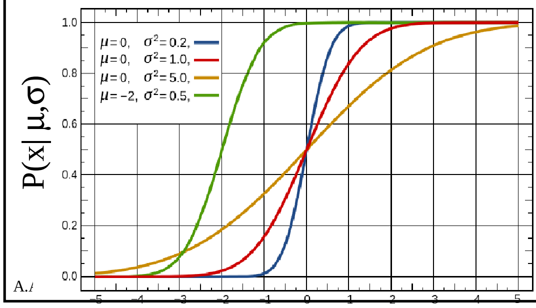
Caso continuo: i valori che X può assumere sono infiniti. Devo trovare un modo per definirne la probabilità. Descrizione **analitica** mediante la funzione densità di probabilità. Si considera la probabilità che x cada in un certo intervallo.

Valgono le stesse relazioni del caso discreto, dove alla somma si sostituisce l'integrale.

$$P(X = x \in [\bar{x}, \bar{x} + \Delta x]) \int_{\bar{x}}^{\bar{x}+\Delta x} \int_{-\infty}^{+\infty} p(x, y)dxdy$$

Distribuzioni notevoli: la Gaussiana





$$p(x|\mu,\sigma) = \frac{1}{\left(\sqrt{2\pi\Sigma^2}\right)^D} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\Sigma}\right)^2\right]$$

D = dimensione, in questo caso D = 1

Pr(| X-μ | < σ) = 0.68268

Pr(| X-μ | < 2σ) = 0.95452

Pr(| X-μ | < 3σ) = 0.9973

I momenti di una variabile statistica

$$\mu^k(X) = \int_{-\infty}^{+\infty} (x-a)^k \, p(x) dx$$ Momento rispetto ad a, solitamente alla media

$$E[X] = \int_{-\infty}^{+\infty} x \, p(x)$$ Valore atteso (Expeted value) di X = media distribuzione

$$E\left[(X-\mu)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2 \, p(x)$$ Varianza (σ^2)

$$E\left[(X-\mu)^3\right] = \int_{-\infty}^{+\infty} (x-\mu)^3 \, p(x)$$ Asimmetria

$$E\left[(X-\mu)^4\right] = \int_{-\infty}^{+\infty} (x-\mu)^4 \, p(x)$$ Kurtosi – peso delle code di p(x)

Overview

Esempi sul teorema di Bayes

Densità di probabilità

Stima alla massima verosimiglianza