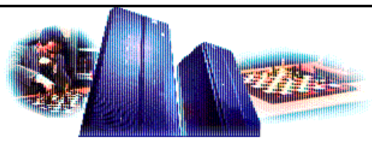
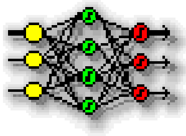




Sistemi Intelligenti
Introduzione all'inferenza statistica

Alberto Borghese

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

Overview

Teoria della probabilità e teorema di Bayes

Esercizi

Probabilità

$$P(A=a_i)=\lim_{N->\infty}\frac{n_{A=a_i}}{N}=\lim_{N->\infty}\frac{n_i}{N}$$

Per il teorema del limite centrale la frequenza di un evento su infinite realizzazioni è uguale alla sua probabilità.

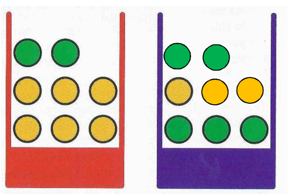
Supponiamo A = {a_1, a_2}
La probabilità che si verifichi uno tra tutti i casi possibili è sempre 1.
Ovverosia la somma delle probabilità di tutti gli eventi (se mutuamente esclusivi) somma 1.

$$P(A=a_1)\cup P(A=a_2)=\lim_{N->\infty}\frac{n_{A=a_1}}{N}+\lim_{N->\infty}\frac{n_{A=a_2}}{N}=\lim_{N->\infty}\frac{n_{A=a_1}+n_{A=a_2}}{N}=1$$

$$P(A)=P(A=a_1)+P(A=a_2)=1$$

Un esempio

Box = {r(ed), b(lue)}
Fruit = {a(pple), o(range)} (apple sono mele verdi)



Supponiamo di prendere un frutto alla volta e di sostituirlo con un frutto dello stesso tipo (sampling with replacement)

Supponiamo di selezionare Box = *r* 40% delle volte e
Box = *b* il 60% delle volte.

P(Box = r) = 0,4 P(Box = b) = 0,6 P(Box) = 1

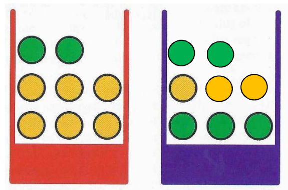
Formulazione statistica

	Red	Blue
Apple		n_{ij}
Orange		

y_j

x_i

$X = \{x_i\} = \{\text{blue, red box}\}$
$Y = \{y_j\} = \{\text{apple, orange}\}$



N trials – z_{ij} is the sample observed in each trial.

n_{ij} is the number of trials in which x_i and y_j were observed.

PROBABILITA' CONGIUNTA

$$\sum_i \sum_j n_{i,j} = N$$

Esempio (N=80)

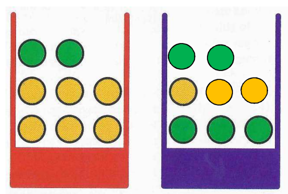
Myself

4/10 (32) → Red Box

6/10 (48) → Blue Box

Red Box: 2/8 (8) → apple; 6/8 (24) → orange

Blue Box: 5/8 (30) → apple; 3/8 (18) → orange



	Red	Blue
Apple	8	30
Orange	24	18

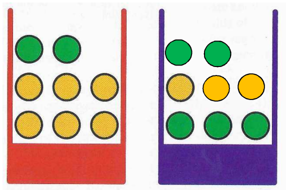
y_j

x_i

$$\sum_{i,j} n_{i,j} = N \quad = 8 + 24 + 30 + 18 = 80$$

Some questions

- What is the probability that the selection procedure will produce an apple?
- What is the probability that the box selected was the blue one if we pick up an orange?



Probabilità congiunta

- What is the probability that the blue box is chosen **and** an apple is selected?

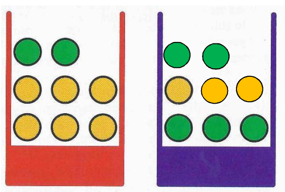
Myself

4/10 (32) 6/10 (48)

Red Box Blue Box

2/8 (8) 6/8 (24) 5/8 (30) 3/8 (18)

apple orange apple orange



	Red	Blue
Apple	**4/10*2/8=8/80 – 8**	**6/10*5/8=30/80 – 30**
Orange	4/10*6/8=24/80 - 24	6/10*3/8=18/80 - 18

y_j

x_i

$$\boxed{P(X=x_i; Y=y_j)} = \lim_{N \to +\infty} \frac{n_{i,j}}{N}$$

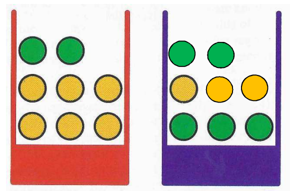
$$\boxed{P(X=b; Y=a)} = 30/80$$

Probabilità condizionata

Myself

4/10 (32) — Red Box
6/10 (48) — Blue Box

Red Box: 2/8 (8) apple; 6/8 (24) orange
Blue Box: 5/8 (30) apple; 3/8 (18) orange



	Red	Blue
Apple	4/10*2/8=8/80 – 8	6/10*5/8=30/80 – 30
Orange	4/10*6/8=24/80 - 24	6/10*3/8=18/80 - 18

y_j

x_i

$$\boxed{P(Y=y_j \mid X = x_i)} = \lim_{N \to +\infty} \frac{n_{ij}}{\sum_j n_{ij}}$$

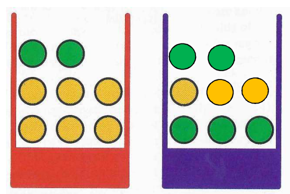
$$\boxed{P(Y=a \mid X = b)} = \frac{n_{a,b}}{n_{a,b} + n_{o,b}} = \frac{30}{30+18} = 5/8$$

Probabilità semplice

Myself

4/10 (32) — Red Box
6/10 (48) — Blue Box

Red Box: 2/8 (8) apple; 6/8 (24) orange
Blue Box: 5/8 (30) apple; 3/8 (18) orange



	Red	Blue
Apple	4/10*2/8=8/80 – 8	6/10*5/8=30/80 – 30
Orange	4/10*6/8=24/80 - 24	6/10*3/8=18/80 - 18

y_j

x_i

$$\boxed{P(X = r)} = \lim_{N \to +\infty} \frac{\sum_j n_{rj}}{N} = \frac{8}{80} + \frac{24}{80} = \frac{32}{80} = 4/10$$

$$\boxed{P(Y = o)} = \lim_{N \to +\infty} \frac{\sum_i n_{i,o}}{N} = \frac{24}{80} + \frac{18}{80} = \frac{42}{80}$$

$$\boxed{P(Y = a)} = \lim_{N \to +\infty} \frac{\sum_i n_{i,a}}{N} = \frac{8}{80} + \frac{30}{80} = \frac{38}{80}$$

Probabilità marginale

$$\boxed{P(X = r)} = \lim_{N \to +\infty} \frac{\sum_j n_{rj}}{N} = \frac{8}{80} + \frac{24}{80} = \frac{32}{80} = 4/10$$

E' la probabilità ottenuta sommando le probabilità delle altre variabili (mettendo ai margini), in questo caso la variabile Y (frutto)

$$\boxed{P(X = r)} = \lim_{N \to +\infty} \frac{\sum_j n_{r,j}}{N} = \sum_j P(X = r, Y = y_j)$$

$$\boxed{P(Y = o)} = \lim_{N \to +\infty} \frac{\sum_i n_{i,o}}{N} = \sum_j P(X = x_i, Y = o)$$

	Red	Blue
Apple	4/10*2/8=8/80 – 8	6/10*5/8=30/80 – 30
Orange	4/10*6/8=24/80 - 24	6/10*3/8=18/80 - 18

y_j

x_i

Riassumendo

$$\boxed{P(X=x_i; Y=y_j)} = \lim_{N \to +\infty} \frac{n_{i,j}}{N}$$

Probabilità congiunta

$$\boxed{P(X=x_i)} = \lim_{N \to +\infty} \frac{\sum_j n_{ij}}{N}$$

Probabilità semplice (marginale)

$$\boxed{P(Y=y_j \mid X = x_i)} = \lim_{N \to +\infty} \frac{n_{ij}}{\sum_j n_{ij}}$$

Probabilità condizionata

$$\boxed{P(X=x_i; Y=y_j)} = \frac{n_{i,j}}{N} = \frac{n_{ij}}{\sum_j n_{ij}} \frac{\sum_j n_{ij}}{N} = \boxed{P(Y=y_j \mid X = x_i)} \boxed{P(X=x_i)}$$

$$\boxed{P(X=x_i; Y=y_j)} = \frac{n_{ij}}{\sum_i n_{ij}} \frac{\sum_i n_{ij}}{N} = \boxed{P(X=x_i \mid Y = y_j)} \boxed{P(Y=y_j)}$$

In forma più compatta

$\boxed{P(X=x_i; Y=y_j) = P(X,Y)}$ Probabilità congiunta

$\boxed{P(X=x_i) = P(X)}$ Probabilità semplice (marginale)

$\boxed{P(Y=y_j \mid X = x_i) = P(Y|X)}$ Probabilità condizionata

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

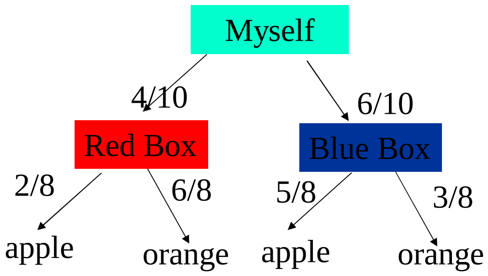
$\boxed{P(X,Y) = P(X|Y)P(Y)}$ $\boxed{P(Y) = \sum_X P(X,Y)}$

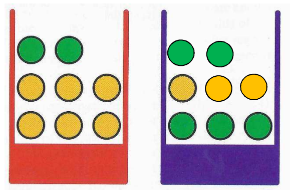
Teorema di Bayes

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

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

Il teorema di Bayes nel nostro caso





Supponiamo di conoscere P(X), probabilità di scelta del box, e la P(Y|X), probabilità di avere una mela (arancia) se scegliamo un certo box, possiamo determinare la probabilità assoluta (semplice) di scegliere un certo frutto, P(Y)?

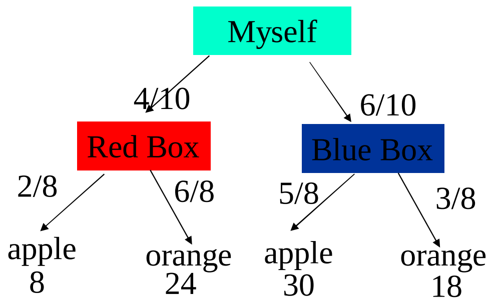
Supponiamo di non conoscere P(X), probabilità di scelta del box, conosciamo la probabilità P(Y|X) e P(Y). Possiamo determinare P(X)?

Determino P(Y)

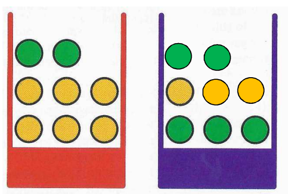
P(Y=apple) ≠ (2+5) / 16 = 7/16
P(Y=orange) ≠ (6+3)/16 = 9/16

P(Y=apple | X = blue) = 5/8
P(Y=orange | X = blue) = 3/8

P(Y=apple | X = red) = 2/8
P(Y=orange | X = red) = 6/8



P(Y=apple | X = blue) + P(Y=orange | X = blue) = 1
P(Y=apple | X = red) + P(Y=orange | X = red) = 1



P(Y=apple) = P(Y=apple | X = blue) P(X=blue) +
P(Y=apple | X = red) P(X=red) = 5/8*6/10 + 2/8*4/10 = 38/80 ≠ 7/16
P(Y=orange) = P(Y=orange | X = blue) P(X=blue) +
P(Y=orange | X = red) P(X=red) = 3/8*6/10 + 6/8*4/10 = 42/80 ≠ 9/16

Determino P(Y) - tabella

	Red	Blue
Apple	4/10*2/8	6/10*5/8
Orange	4/10*6/8	6/10*3/8

y_j

x_i

Sommo lungo x, marginalizzo x.

P(Y=apple) = 5/8*6/10 + 2/8*4/10 = 38/80
P(Y=orange) = 3/8*6/10 + 6/8*4/10 = 42/80

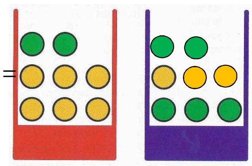
P(X=red) = 4/10*2/8 + 4/10*6/8 = 4/10
P(X=blue) = 6/10*5/8 + 6/10*3/8 = 6/10

Determino P(X|Y)

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

Myself
4/10 6/10
Red Box Blue Box
2/8 6/8 5/8 3/8
apple 8 orange 24 apple 30 orange 18

P(X=red | Y=orange) = P(Y=orange|X=red) P(X=red)/P(Y=orange) =
(6/8*4/10) / (21/40) = 24/42 > 4/10
P(X=blue | Y=orange) = P(Y=orange|X=blue) P(X=blue)/P(Y=orange) =
(3/8*6/10) / (21/40) = 18/42 < 6/10

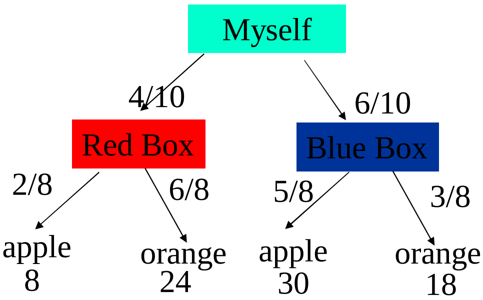


P(X=red | Y=apple) = P(Y=apple | X=red) P(X=red) / P(Y=apple) =
(2/8*4/10) / (19/40) = 8/38 << 4/10
P(X=blue | Y=apple) = P(Y=apple | X=blue) P(X=blue) / P(Y=apple) =
(5/8*6/10) / (19/40) = 30/38 >> 6/10

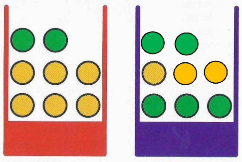
Interpretazione

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

P(X=red | Y=apple) = 8/38 << 4/10
P(X=blue | Y=apple) = 30/38 >> 6/10



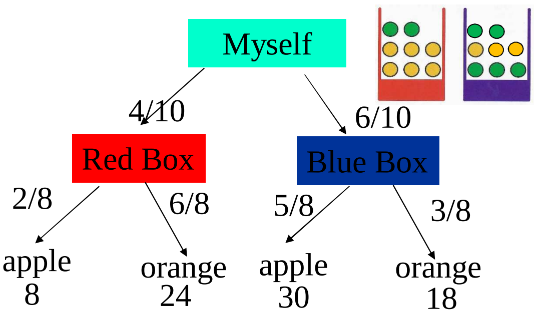
Correggo la probabilità a-priori, P(X) con le informazioni raccolte, P(Y), e ottengo una nuova valutazione della probabilità di X (che dipende da Y), P(X | Y) detta probabilità a-posteriori.



	Red	Blue	
Apple	4/10*2/8=8/80 – 8	6/10*5/8=30/80 – 30	y_j
Orange	4/10*6/8=24/80 - 24	6/10*3/8=18/80 - 18	
		x_i	

Importanza

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



Questo è un tipico esempio di problema **inverso**.

Raccogliamo delle misure Y e vogliamo determinare da quale sistema (modello probabilistico) possono essere state generate.

Possiamo inserire delle informazioni statistiche (a-priori) su X, cioè sulla forma del modello (e.g. smoothness)

Overview

Teoria della probabilità e teorema di Bayes

Esercizi

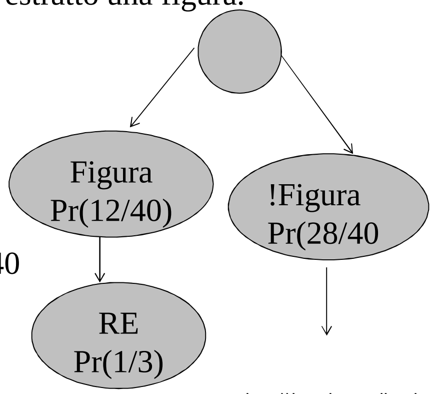
Probabilità condizionata

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

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

$P(Y \mid X) = 1/3$

$P(Y) = P(Y/X)\, P(X) = 1/3 \;12/40 = 4/40$



Esempio I

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

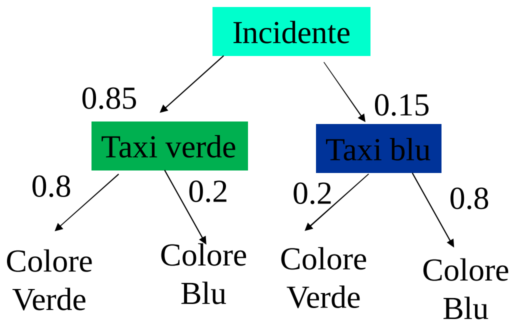
Succede un incidente in cui è coinvolto un taxi.
Un testimone dichiara che il taxi era blu. Era sera e l'affidabilità del testimone è stata valutata dell'80%.

Qual è la probabilità che il taxi fosse effettivamente blu?

Non è l'80%!

Esempio - I

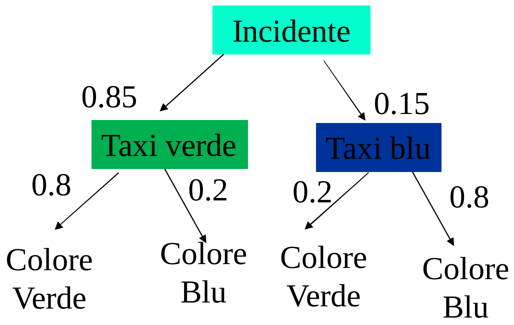
$X = \{T_{blu}, T_{verde}\}$
$Y = \{C_{blu}, C_{verde}\}$



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

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

Esempio - I



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

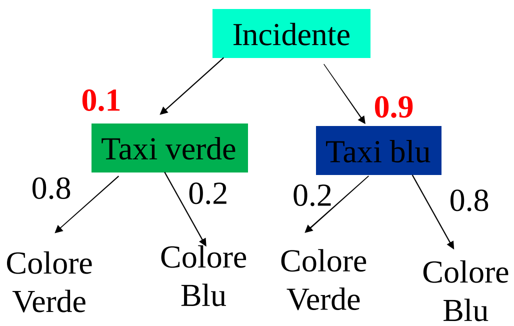
$P(X=T_{blu}|Y_{blu}) =$
$P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu})$

$P(Y_{blu})$ = Probabilità marginale di Y (probabilità semplice) =

$P(Y_{blu}|X_{blu})P(X_{blu}) + P(Y_{blu}|X_{verde})P(X_{verde}) = 0.8*0.15 + 0.2*0.85 = 0.29$

$P(X=T_{blu}|Y_{blu}) = P(Y_{blu}|X_{blu})P(X_{blu}) / P(Y_{blu}) = 0.8*0.15 / 0.29 = 0.41$

Esempio - I



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

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

$P(Y_{blu})$ = Probabilità marginale di Y (probabilità semplice) =

$P(Y_{blu}|X_{blu})P(X_{blu}) + P(Y_{blu}|X_{verde})P(X_{verde}) = 0.8*0.9 + 0.2*0.1 = 0.74$

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

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?

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%:

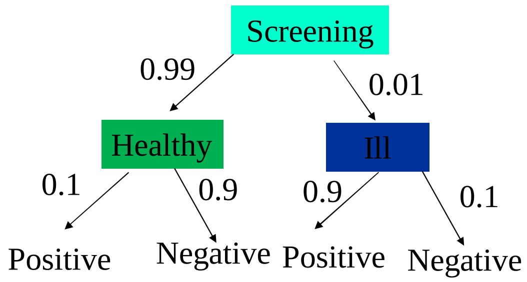
$$\text{sensitività} = \frac{n_{positive}}{N_{ill}} \qquad \Rightarrow P(Y=positive \mid X=ill)$$

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

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

Esempio II

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



P(X=Ill | Y=Positive) = P(Y=Positive | X=Ill)P(X=Ill) / P(Y=Positive)

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*0.1 = 0.108

Esempio - II

P(Y=Positive = P(Y=Positive | X=Ill) + P(Y=Positive | X=Healthy) =
0.009 + 0.99*0.1 = 0.108

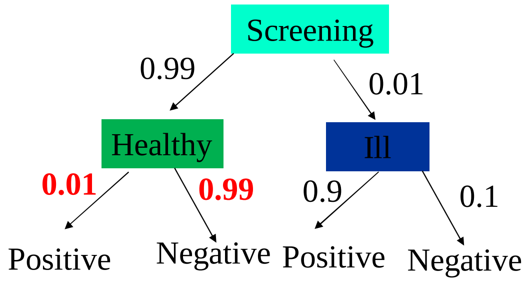
10.8% di probabilità di avere un esame positivo a fronte di uno 0.01% di donne malate! Solo lo 0,9% proviene da donne effettivamente malate, le altre sono false positive

Infatti 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}



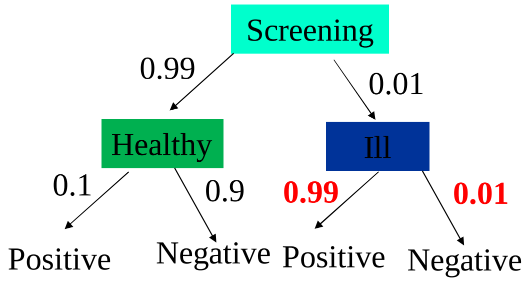
P(Y=Positive | X=Ill) = 0.9 * 0.01 = 0.009

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, X, a partire dalla probabilità associata alla frequenza di una certa misura, P(Y), dalla probabilità che lega la misura all'evento, associata alla frequenza relativa dell'evento associato alla misura, P(Y|X), e dalla probabilità nota a-priori, P(X), dell'evento.

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

Viene utilizzata nei problemi inversi.

Overview

Teoria della probabilità e teorema di Bayes

Esercizi