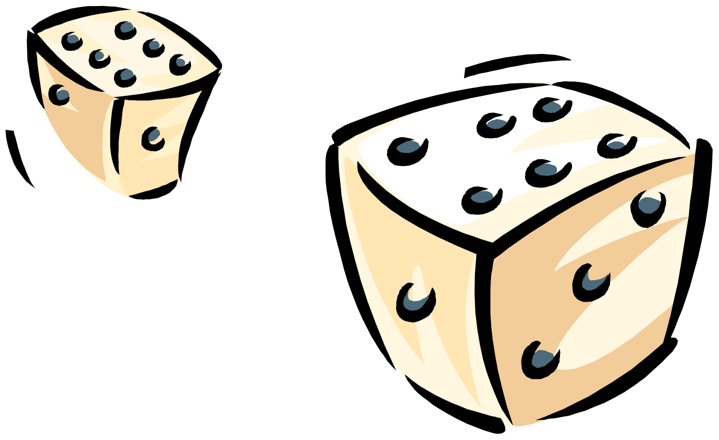
Cenni di calcolo delle probabilità

Quando si compie un esperimento o una serie di prove i possibili risultati osservati prendono il nome di eventi.

Se tali risultati non sono conosciuti a priori allora si definiscono eventi aleatori.



Proprietà elementari delle probabilità

la probabilità di un evento E che si indica con P(E)

- è un numero sempre positivo $P(E) \geq 0$
- compreso tra 0 ed 1: $0 \leq P(E) \leq 1$

- se E è un evento certo allora la probabilità di E è $P(E)=1$

- se E è un evento impossibile $P(E)=0$

- dato un processo sperimentale che genera n risultati (eventi) disgiunti, $A_1, A_2, A_3, \ldots, A_n$, la probabilità di un evento A_i è un numero non negativo: $P(A_i) \geq 0$

- la somma delle probabilità di tutti gli eventi possibili mutuamente esclusivi è uguale ad 1: $P(A_1)+P(A_2)+\ldots+P(A_n)=1$

Dati due eventi A, B, la probabilità che si verifichi l'evento A o l'evento B o entrambi è :

$$P(B \cup A) = P(A) + P(B) - P(B \cap A)$$

Se i due eventi A e B sono mutuamente esclusivi allora

$$P(B \cup A) = P(A) + P(B)$$

legge della somma.

Talvolta tutti i possibili risultati possono essere un sottoinsieme del totale.
Se A e B sono il risultato di un esperimento può accadere che il verificarsi dell’evento B sia modificato dal fatto che si sia già verificato l’evento A.
Si dice allora che l’evento B è *condizionato* da A e la probabilità che si verifichi l’evento B è condizionata dalla probabilità dell’evento A:

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$

Se due eventi sono indipendenti allora la probabilità di B∩A è data dal prodotto della probabilità di A per la probabilità di B:

$$P(B \cap A) = P(A) * P(B)$$

legge del prodotto

TEOREMA DI BAYES

- **Il teorema di Bayes viene definito il teorema delle probabilità a posteriori…**

…infatti la conoscenza del realizzarsi di un particolare evento condizionante…

- **SINTOMO**
- **SEGNO**
- **TEST DI LABORATORIO**

…può modificare la probabilità a priori di un altro evento…

- **MALATTIA**

…aumentandone o riducendone il valore iniziale.

Applicazioni del teorema di Bayes in medicina

<u>La formulazione di diagnosi</u> :

in questo caso l'evento A1, A2, A3, …, Ak, sono tutte le possibili diagnosi, mutuamente esclusive, che corrispondono ad un complesso di sintomi, evento B.

E' indispensabile conoscere a priori:

- **la probabilità di avere quei sintomi B dato che si ha la malattia A1**

- **la probabilità di avere la malattia A1 (prevalenza della patologia).**

Applicazioni del teorema di Bayes in medicina

<u>Gli screenings</u>:

il teorema di Bayes permette di determinare il valore predittivo positivo di un test di screening.

E' indispensabile conoscere a priori:

- **la specificità (capacità di trovare i soggetti veramente negativi tra i sani);**

- **la sensibilità (capacità di indicare come positivi soggetti che sono malati);**

- **la prevalenza della patologia.**

PROBLEMI DI DIAGNOSI

Ipotizziamo che un paziente si presenti in un ambulatorio con astenia e storia di calcoli renali.

Una delle diagnosi possibili è iperparatiroidismo ⇒ Malattia: evento A1
Non malattia: evento A2

Ipotizzando che la prevalenza della patologia sia circa del 2% possiamo ritenere che:

P (A1) = 0,02 = 2% (avere iperparatiroidismo)

P (A2) = 1-0,02 = 0,98 = 98% (non avere iperparatiroidismo)

PROBLEMI DI DIAGNOSI

Applicando un test diagnostico potremmo cercare **conferme** alla diagnosi → Valutazione calcio sierico

Supponiamo di avere un valore di calcemia compatibile con la diagnosi di iperparatiroidismo. → Test di laboratorio positivo: T+
P(T+) = probabilità dell'evento T+

Tale situazione può verificarsi anche con altre diagnosi.

Si può essere
T+ con iperparatiroidismo (A1 e T+)
T+ senza iperparatiroidismo (A2 e T+)

PROBLEMI DI DIAGNOSI

Il problema diagnostico può essere posto in questi termini:
Qual è la probabilità di avere iperparatiroidismo dato che la calcemia è alta?

$$\mathbf{P(A1/T+) = \frac{P(A1\ e\ T+)}{P(T+)}} \qquad \mathbf{(1)}$$

$$\mathbf{P(A1\ e\ T+)=P(A1/T+)\ P(T+)}$$

$$\mathbf{P(A1\ e\ T+)=P(T+/A1)\ P(A1)}$$

PROBLEMI DI DIAGNOSI

$P(T+) = P[(A1 \text{ e } T+) \text{ o } (A2 \text{ e } T+)] =$
$= P(A1 \text{ e } T+) + P(A2 \text{ e } T+)$

$P(A1 \text{ e } T+)=P(T+/A1)\ P(A1)$

Percentuale di individui con calcemia alta dato che hanno la patologia
90%

Prevalenza della patologia
2%

$P(A2 \text{ e } T+)=P(T+/A2)\ P(A2)$

Percentuale di individui con calcemia alta dato che non hanno la patologia
5%

Prevalenza della non-patologia
98%

PROBLEMI DI DIAGNOSI

Sostituendo i dati noti nella (1) avremo:

$$\mathbf{P(A1/T+)} = \frac{\mathbf{P(A1\ e\ T+)}}{\mathbf{P(T+)}} =$$

$$= \frac{\mathbf{P(T+/A1)\ P(A1)}}{\mathbf{[P(T+/A1)\ P(A1)]+[P(T+/A2)\ P(A2)]}} =$$

= (0,9*0,02) / [(0,9*0,02) + (0,05*0,98)]= 0,269 = circa 27%

PROBLEMI DI DIAGNOSI

La probabilità di avere iperparatiroidismo è passata dal 2%
Senza l’informazione relativa all’ipercalcemia…

PROBABILITA’ A PRIORI

…ad un valore del 27% dopo che la calcemia è risultata elevata

PROBABILITA’ A POSTERIORI

SCREENING

Il teorema di Bayes consente di determinare il valore predittivo positivo di un test di screening

ESEMPIO

Si vuole iniziare una campagna di screening per la individuazione di soggetti positivi alla tubercolina

La sensibilità del test è 96%
La specificità del test è 94%
La prevalenza di TBC nella comunità è del 1%
La popolazione su cui fare lo screening è di 10000 individui

SCREENING

	Malati	Sani	Totale
Tubercolino +	(a) 96	(b) 594	690
Tubercolino -	(c) 4	(d) 9306	9310
Totale	100	9900	10000

$$Sensibilit\grave{a} = \frac{a}{a+c}$$ $\boxed{\mathbf{P(T+/M)}}$ **96%**

$$Specificit\grave{a} = \frac{d}{b+d}$$ $\boxed{\mathbf{P(T-/S)}}$ **94%**

SCREENING

	Malati	Sani	Totale
Tubercolino +	(a) 96	(b) 594	690
Tubercolino -	(c) 4	(d) 9306	9310
Totale	100	9900	10000

$$Falsi + = \frac{b}{b+d}$$

P(T+/S)

100%-94%=6%

$$Falsi - = \frac{c}{a+c}$$

P(T-/M)

100%-96%=4%

SCREENING

	Malati	Sani	Totale
Tubercolino +	(a) 96	(b) 594	690
Tubercolino -	(c) 4	(d) 9306	9310
Totale	100	9900	10000

$$valore\_predittivo\,(+) = \frac{a}{a+b}$$

$$\boxed{\mathbf{P(M/T+)}} = \frac{\mathbf{P(T+/M)\ P(M)}}{\mathbf{[P(T+/M)\ P(M)]+[P(T+/S)\ P(S)]}}$$

SCREENING

$$\boxed{P(M/T+)} = \frac{P(T+/M)\,P(M)}{[P(T+/M)\,P(M)]+[P(T+/S)\,P(S)]} =$$

$$= \frac{\text{Sensibilità x prevalenza}}{\text{(Sensibilità x prevalenza) + [(Falsi positivi) x (1-prevalenza)]}}$$

$$= \frac{0{,}96 \times 0{,}01}{(0{,}96 \times 0{,}01) + (0{,}06 \times 0{,}99)} = \mathbf{0{,}139} \Rightarrow \mathbf{13{,}9\%}$$