Appunti della Lezione 3

Teorema di Bayes

Spesso, l'osservazione di un dato sperimentale, che chiameremo Evento E, e' condizionata da una diversa situazione iniziale, che chiameremo ipotesi H_i. Per fare un esempio concreto, possiamo immaginare l'osservazione di un evento "$E =$ fulmine", che in qualche modo dipende dalla direzione del vento (Ipotesi "$H_i =$ maestrale, libeccio, bora, ...). Supponiamo che una qualsiasi ipotesi escluda tutte le altre e che la somma di tutte le ipotesi riempia lo spazio totale delle ipotesi (la *rosa dei venti*), ovvero che

$$\textstyle\sum_i H_i = S, \qquad H_i H_j = 0, \qquad \forall i,j$$

Questo ci permette di scrivere la seguente importante relazione, che, nel nostro esempio, ci da' la probabilita' totale di osservare un fulmine:

$$\begin{aligned} P(E) &= P(E+S) \\ &= P(E+\textstyle\sum H_i) \\ &= P(EH_1) + P(EH_2) + ... \\ &= P(E|H_1)P(H_1) + P(E|H_2)P(H_2) + ... \\ &= \textstyle\sum P(E|H_i)P(H_i) \end{aligned}$$

che e' anche detta formula delle probabilita' totali (o di partizione dell'evento certo). [1] In pratica noi misuriamo la probabilita' $P(E|H_i)$ di vedere un fulmine in presenza di un certo tipo di vento e, conoscendo le probabilita' $P(H_i)$, siamo in grado di ottenere il valore della probabilita' totale dell'evento "fulmine".

Il teorema di Bayes esprime una relazione importante tra queste quantita'

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum P(E|H_i)P(H_i)}$$

E' facilmente dimostrabile, partendo dalla relazione $P(A|B)P(B) = P(BA) = P(AB) = P(B|A)P(A)$ ed utilizzando la formula delle probabilita' totali

$$\begin{aligned} P(H_i|E)P(E) &= P(E|H_i)P(H_i) \\ P(H_i|E) &= \frac{P(E|H_i)P(H_i)}{P(E)} \\ &= \frac{P(E|H_i)P(H_i)}{\sum P(E|H_i)P(H_i)} \; c.v.d. \end{aligned}$$

[1] Un caso particolarmente interessante della formula delle probabilita' totali, utilizzato in molte applicazioni, e' il caso in cui esiste un'unica ipotesi H e la sua negazione $\overline{H}$. In questo caso l'espressione diventa

$$P(E) = P(E|H)P(H) + P(E|\overline{H})P(\overline{H})$$

Questo teorema puo' essere utilizzato in due modi diversi, con un approccio di tipo "Bayesiano" quando non sono note a priori le probabilita' $P(H_i)$, o con un approccio di tipo frequentista quando si assumono date le probabilita' $P(H_i)$. In entrambi in casi, le probabilita' condizionate provengono da una misura. L'approccio che noi utilizziamo e' di tipo frequentista.

Vediamo ora un esempio. La radiazione cosmica e' composta per il 90% di muoni μ ("elettroni pesanti"), e per il 10% di elettroni (questa suddivisione e' arbitraria, ma molto vicina alla realta'). Consideriamo queste come le nostre due ipotesi iniziali. Esse soddisfano la condizione che siano mutuamente escludenti e che coprano lo spazio delle ipotesi, per cui $P(e) + P(\mu) = 1.00$. Supponiamo di avere costruito uno strumento per rivelare queste particelle e per distinguerle, con un'efficienza migliore nel rivelare gli elettroni ($P(R|e) = 0.90$ e $P(R|\mu) = 0.30$). Vogliamo determinare qual e' la percentuale di eventi di elettroni misurati sul totale degli eventi rivelati. Possiamo applicare il teorema di Bayes, avendo i seguenti dati

$$
\begin{aligned}
P(e) &= 0.10 \\
P(\mu) &= 0.90 \\
P(e) + P(\mu) &= 1.0 \quad (certezza) \\
P(R|e) &= 0.90 \\
P(\overline{R}|e) &= 1 - P(R|e) = 0.10 \\
P(R|\mu) &= 0.30 \\
P(\overline{R}|\mu) &= 1 - P(R|\mu) = 0.70 \\
P(R) &= P(R|e)P(e) + P(R|\mu)P(\mu) = 0.36 \\
P(e|R) &= \frac{P(R|e)P(e)}{P(R)} = 0.25
\end{aligned}
$$

Che ci dice che degli eventi osservati, il 25% sono eventi di elettroni; in questo modo, grazie al diverso comportamento dei rivelatori, la percentuale di elettroni osservati e' superiore a quella degli eventi originali della radiazione cosmica. Possiamo a questo punto aumentare la percentuale degli eventi osservati, introducendo un secondo rivelatore che metteremo in coincidenza con il primo. Dobbiamo pertanto calcolare quantita' iniziali diverse da prima, in particolare dobbiamo calcolare qual e' la probabilita' che entrambi i rivelatori funzionino in presenza di elettroni e di muoni.

$$
\begin{aligned}
P(RR|e) &= P(R|e)P(R|e) = 0.90^2 = 0.81 \\
P(\overline{RR}|e) &= 1 - P(R|e) = 0.10 \\
P(RR|\mu) &= P(R|\mu)P(R|\mu) = 0.30^2 = 0.09 \\
P(\overline{RR}|\mu) &= 1 - P(R|\mu) = 0.81 \\
P(RR) &= P(RR|e)P(e) + P(RR|\mu)P(\mu) = 0.162 \\
P(e|RR) &= \frac{P(RR|e)P(e)}{P(RR)} = 0.50
\end{aligned}
$$

Si vede come la richiesta che entrambi i rivelatori diano un segnale porta ad un risultato molto diverso dal dato di partenza, ovvero che la meta' degli eventi registrati sono associati alla presenza di un elettrone rispetto alla percentuale intrinseca della radiazione cosmica.

Va notato innanzitutto che la probabilita‘ totale $P(RR)$ di rivelare una particella con due rivelatori NON e‘ uguale al prodotto delle probabilita‘ ($P(RR) \neq P(R)P(R)$) in quanto questa e‘ ricavata da probabilita‘ condizionate, ed e‘ facile verificare che il segno “=” vale nel caso in cui $P(R|e) = P(R|\mu) = P(R)$, ovvero nel caso in cui la probabilita‘ di rivelare una particella sia indipendente dalla stessa.

Possiamo ribaltare il problema, ovvero chiederci come costruire i nostri rivelatori, per ottenere il risultato che il 99% delle particelle rivelate siano elettroni $P(e|RR) = 0.99$. In questo caso, abbiamo due quantita‘ che devono essere calcolate per avere il risultato richiesto, ovvero $P(RR|e)$ e $P(RR|\mu)$. Queste due quantita‘ saranno legate da una relazione (lineare) che possiamo calcolare con alcuni passaggi

$$P(e|RR) = [P(RR|e)P(e)]/[P(RR|e)P(e) + P(RR|\mu)P(\mu)]$$

$$P(e|RR)(P(RR|e)P(e) + P(RR|\mu)P(\mu)) = P(RR|e)P(e)$$
$$P(RR|e)P(e)[1 - P(e|RR)] = P(RR|\mu)P(e|RR)P(\mu)$$
$$P(RR|e) = P(RR|\mu)[P(e|RR)P(\mu)]/[(1 - P(e|RR))P(e)]$$

e sostituendo i numeri otteniamo

$$\begin{aligned} P(RR|e) &= P(RR|\mu)[0.99 \cdot 0.9]/[(1 - 0.99) \cdot 0.1] \\ &= P(RR|\mu) \cdot 891 \end{aligned}$$

questo significa costruire un sistema di rivelatori che sia circa mille volte piu‘ efficiente nel rivelare elettroni che muoni. Se pensiamo di mettere due rivelatori con le stesse caratteristiche, questo significa costruire un rivelatore che sia trenta volte piu’ efficiente nel rivelare elettroni che muoni.