

Esame di TEORIA DEI SEGNALI

Quesito A53

Il sig. Rossi teme di avere una certa malattia, pertanto si sottopone a un test diagnostico che ha le seguenti caratteristiche:

$$P(R^c | M) = 10^{-2} \quad (\text{Prob. di falso negativo})$$

$$P(R | M^c) = 6 \cdot 10^{-2} \quad (\text{Prob. di falso positivo})$$

dove R ed M sono i seguenti eventi:

$R = \{\text{Il test rivela presenza di malattia}\}$ (ossia il test è positivo)

$M = \{\text{La malattia è presente}\}.$

Il test del sig. Rossi risulta positivo, ma il medico curante lo rassicura: "Sì, il test è positivo, ma nonostante ciò lei ha l'80% di probabilità di essere sano".

2a) Si calcoli la probabilità che un individuo scelto a caso fra la popolazione sia affetto dalla malattia (ovvero l'incidenza della malattia sulla popolazione).

[Facoltativo: alla luce del risultato si commenti l'affermazione del medico, apparentemente paradossale.]

2b) Per sicurezza il test viene ripetuto (in modo indipendente) e il risultato è nuovamente positivo. Qual è la probabilità che il sig. Rossi sia effettivamente malato dato il doppio test positivo?

Quesito A53 (soluzione)

Dal testo si ricavano direttamente le seguenti probabilità:

$$P(\bar{R} | M) = 10^{-2} \quad \text{da cui: } P(R | M) = 1 - P(\bar{R} | M) = 0,99$$

$$P(R | \bar{M}) = 6 \cdot 10^{-2} \quad \text{da cui: } P(\bar{R} | \bar{M}) = 1 - P(R | \bar{M}) = 0,94$$

Dalle parole del medico si ricava la probabilità che una persona qualsiasi (il sig. Rossi) non abbia la malattia (sia sano) nel caso in cui il test sia risultato positivo, ossia la seguente probabilità condizionata:

$$P(\bar{M} | R) = 0,8 \quad \text{da cui } P(M | R) = 1 - P(\bar{M} | R) = 0,2 \quad (1)$$

2a) Con questi dati si chiede di calcolare la probabilità che la malattia sia presente in una persona qualsiasi ossia di calcolare $P(M)$.

Come sempre i modi di risolvere il problema possono essere più di uno - Un modo è il seguente:

Applicando la formula di Bayes alla (1) si ha.

$$P(M/R) = \frac{P(R|M) \cdot P(M)}{P(R)} = 0,2 \quad (2)$$

Poiché $P(R|M) = 0,99$ è data, si tratta di esprimere $P(R)$ in funzione di $P(M)$ e di ricavare quest'ultima dalla (2). Dal teorema delle probabilità totali si ha:

$$\begin{aligned} P(R) &= P(R|M) \cdot P(M) + P(R|\bar{M}) [1 - P(M)] = \\ &= 0,99 \cdot P(M) + 6 \cdot 10^{-2} \cdot [1 - P(M)] \simeq 0,074 \end{aligned}$$

Quindi dalla (2):

$$\frac{0,99 \cdot P(M)}{0,99 \cdot P(M) + 6 \cdot 10^{-2} [1 - P(M)]} = 0,2$$

$$0,99 \cdot P(M) = 0,2 \cdot 0,99 \cdot P(M) + 0,2 \cdot 6 \cdot 10^{-2} - 0,2 \cdot 6 \cdot 10^{-2} \cdot P(M)$$

$$\boxed{P(M) = \frac{0,2 \cdot 6 \cdot 10^{-2} \leftarrow 0,012}{0,99 - 0,2 \cdot 0,99 + 0,2 \cdot 6 \cdot 10^{-2} \leftarrow 0,804}} \simeq 0,015 = \boxed{1,5 \cdot 10^{-2}} = \frac{1}{67}$$

Commento

Il risultato paradossale (ovvero anti-intuitivo, apparentemente strano) è che pur essendo il test molto affidabile (prob. di falso positivo e falso negativo molto basse) quando esso rivela malattia (test positivo), nella grande maggioranza dei casi (80%) sbaglia.

Alla luce del risultato trovato risulterà che ciò è dovuto alla bassa incidenza della malattia sulla popolazione ($P(M) = 1,5 \cdot 10^{-2} = 1,5\%$) molto più bassa della probabilità di falso positivo ($P(R|\bar{M}) = 6 \cdot 10^{-2} = 6\%$). Al limite se la malattia non esistesse ($P(M) = 0$) l'80% di errori diventerebbe il 100% infatti tutti i test positivi (che sarebbero il 6% dei test eseguiti) sarebbero dovuti a falsi positivi.

25) Si cerca ora la probabilità $P(M|R_1R_2)$ essendo

$R_1R_2 = \{ \text{Risultato positivo in due esecuzioni indipendenti del test sulla stessa persona} \}.$

Quindi (da Bayes):

$$P(M|R_1R_2) = \frac{P(R_1R_2|M) \cdot P(M)}{P(R_1R_2)}$$

dove (Probs. totali):

$$P(R_1R_2) = P(R_1R_2|M) \cdot P(M) + P(R_1R_2|\bar{M}) \cdot P(\bar{M})$$

Per l'indipendenza si ha:

$$P(R_1R_2|M) = P(R_1|M) \cdot P(R_2|M) = [P(R|M)]^2 = 0,99^2$$

e analogamente:

$$P(R_1R_2|\bar{M}) = [P(R|\bar{M})]^2 = (6 \cdot 10^{-2})^2$$

Quindi:

$$P(R_1 R_2) = (0,99)^2 \cdot 0,015 + (6 \cdot 10^{-2})^2 [1 - 0,015] \simeq 0,0182$$

ed infine

$$P(M | R_1 R_2) = \frac{(0,99)^2 \cdot 0,015}{0,0182} \simeq 0,801 \simeq 80\%$$

Con il secondo test positivo la probabilità di essere malato dello sfortunato org. Romi passa dal 20% all' 80% -

Di fatto il test ripetuto equivale ad un unico test con probabilità di falso positivo:

$$P(R_1 R_2 | \bar{M}) = [P(R | \bar{M})]^2 = (6 \cdot 10^{-2})^2 = 3,6 \cdot 10^{-5}$$

molto minore di $P(M) = 1,5 \cdot 10^{-2}$, quindi affidabile anche con incidenza così bassa -