

A4 Test gravidanza

Test clinici³⁶

Alcuni test clinici, come per esempio quello di gravidanza, danno una risposta che può essere positiva o negativa: nell'esempio, test positivo significa che la donna è incinta, test negativo significa che la donna non è incinta.

Qualche volta il test dà risultati sbagliati; riportiamo una statistica di un laboratorio di analisi, relativamente a 3000 test di gravidanza effettuati:

- fra i 1.000 test risultati negativi, 80 erano sbagliati (cioè la donna era incinta, mentre il test stabiliva che non lo era);
- fra i 2.000 test risultati positivi solo 5 erano sbagliati (cioè la donna non era incinta, mentre il test stabiliva che lo era).

A partire da questi dati risolvere i seguenti quesiti:

- calcolare la probabilità q che un test positivo indichi la gravidanza (cioè determinare *il valore predittivo del test positivo*),
- calcolare la probabilità r che un test negativo escluda la gravidanza (cioè determinare *il valore predittivo del test negativo*),
- calcolare la probabilità p che il test dia indicazioni esatte (cioè determinare *l'efficienza del test*).

Un primo approccio risolutivo

Per rispondere al quesito si può seguire una via puramente numerica.

Riteniamo però più significativo dal punto di vista didattico seguire anche altri due approcci che prevedono il ricorso ad opportune rappresentazioni grafiche. Li illustreremo esaminando la domanda c.

- Il testo afferma che, dei 1000 test positivi ("casi possibili"), solo 5 test sono errati. Quindi sono 1995 gli esiti che indicano correttamente la gravidanza ("casi favorevoli"). La probabilità q richiesta è pertanto:

$$q = \frac{1995}{2000} \cong 0,998$$

- Ragionando in modo analogo, si determina facilmente la probabilità r che un test negativo escluda la gravidanza:

$$r = \frac{920}{1000} = 0,920$$

³⁶ Tratto da Castelnuovo - Gori Giorgi - Valenti, Matematica oggi 2, ed. La Nuova Italia, 1992, pag. 821 n. 214.

Modellizzazione mediante tabella a doppia entrata

Possiamo **suddividere** la popolazione in esame in quattro sottoinsiemi, a seconda dell'esito del test e della presenza o meno di gravidanza³⁷ e rappresentare in modo espressivo la situazione mediante una tabella come la seguente:

	negativi	positivi
no gravidanze	Diagnosi corretta	Falso Positivo: test positivo e no gravidanza
gravidanze	Falso Negativo: test negativo e gravidanza	Diagnosi corretta

Riportiamo sulla tabella i dati forniti nel testo e a partire da essi ricaviamo presto gli altri che servono per rispondere al quesito³⁸ (a destra).

	negativi	positivi	
no gravidanze		5	
gravidanze	80		
Totale	1000	2000	3000

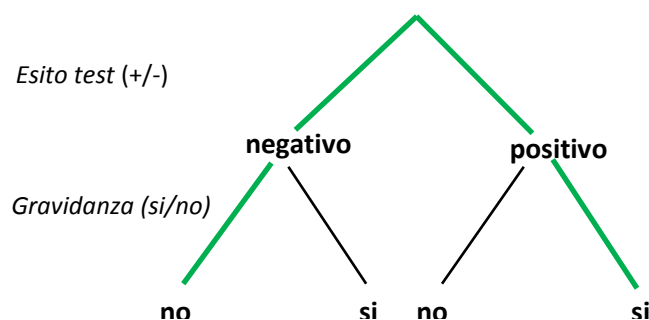
	negativi	positivi	
no gravidanze	920	5	
gravidanze	80	1995	
Totale	1000	2000	3000

Il test fornisce indicazioni esatte per le donne che appartengono ai due sottoinsiemi evidenziati in verde. Pertanto i **"casi favorevoli"** sono $1995 + 920 = 2915$ su una popolazione di 3000 unità, i **"casi possibili"**. In definitiva la probabilità p che il test dia indicazioni esatte è

$$p = \frac{2915}{3000} \approx 0,972.$$

Modellizzazione mediante grafo ad albero

Per comprendere più in profondità la questione può essere utile schematizzare le situazioni che si possono presentare anche mediante un grafo ad albero:



Sono evidenziati in verde i due cammini che rappresentano i casi in cui il test fornisce un esito esatto:

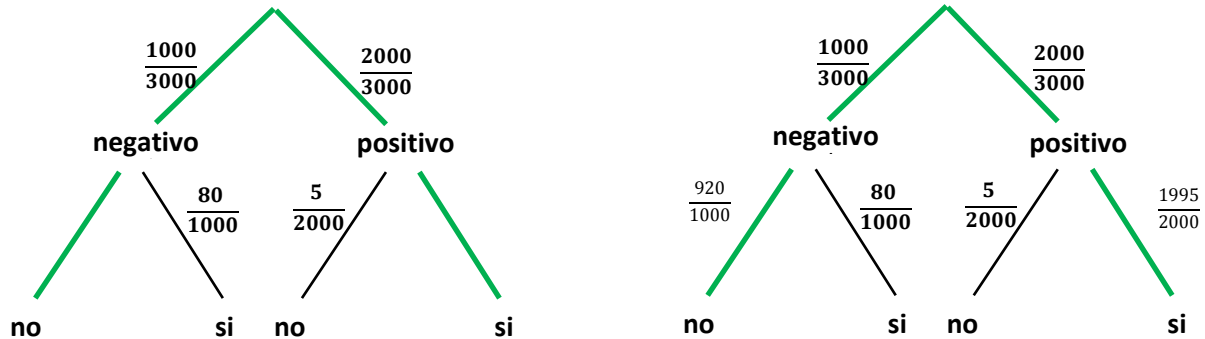
- *test negativo e "no" gravidanza*
- *test positivo e gravidanza*

³⁷ Più precisamente si è ottenuta una *partizione* della popolazione in esame.

³⁸ Il testo afferma che "su 1000 test negativi, 80 erano sbagliati". Questo implica che i restanti 920 test negativi erano esatti, cioè indicavano effettivamente assenza di gravidanza.

Per calcolare le probabilità di questi due eventi, cominciamo con l'indicare accanto ad ogni ramo il valore di probabilità dell'evento "elementare" corrispondente.

Sul grafo posto a sinistra riportiamo le probabilità che si desumono *direttamente* dai dati forniti. Deduciamo³⁹ poi le due probabilità che mancano e le aggiungiamo sul grafo a destra.



La probabilità dell'evento "test negativo e no gravidanza" si può ottenere in modo espressivo **moltiplicando** i valori di probabilità indicati lungo il cammino corrispondente:

$$p(\text{"test negativo e no gravidanza"}) = \frac{1000}{3000} \cdot \frac{920}{1000} = \frac{920}{3000}$$

Analogamente

$$p(\text{"test positivo e gravidanza"}) = \frac{2000}{3000} \cdot \frac{1995}{2000} = \frac{1995}{3000}$$

Il test dà un'indicazione esatta quando accade uno **oppure** l'altro dei due eventi rappresentati in verde. Pertanto la sua probabilità è la **somma** delle probabilità di questi due eventi:

$$p(\text{"test esatto"}) = \frac{920}{3000} + \frac{1995}{3000} = \frac{2915}{3000} \approx 0,972.$$

Osservazioni

- Ciascuna delle quattro celle della tabella (che rappresentano i quattro sottoinsiemi in cui è suddivisa la popolazione) **corrisponde** ad uno dei quattro possibili cammini sul grafo ad albero. Ad esempio alla cella in alto a sinistra nella tabella corrisponde il cammino più a sinistra sul grafo ad albero ("negativo e no gravidanza").
- Per i risultati sottesati all'uso del grafo ad **albero come modello di calcolo** rimandiamo alle osservazioni scritte a commento di altri esercizi (in particolare quello relativo ai *circuiti elettrici*).
- La somma delle probabilità dei rami relativi ad uno stesso nodo vale 1. E ancora 1 vale la somma delle probabilità di tutti i quattro cammini sull'albero:

$$\frac{1995}{3000} + \frac{5}{3000} + \frac{80}{3000} + \frac{920}{3000} = 1$$

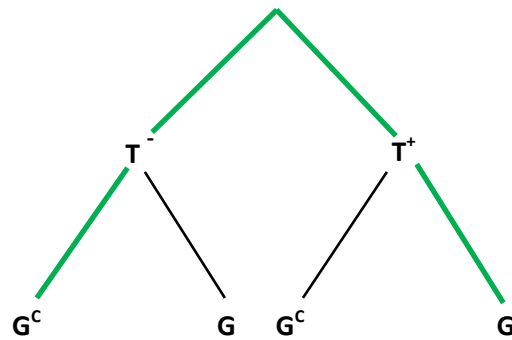
Questi due fatti possono essere utilizzati per controllare i risultati.

- Possiamo esprimere il procedimento in modo conciso, utilizzando le seguenti **notazioni** per indicare gli eventi che intervengono:

$$\begin{aligned} T^+ &= \text{"la donna è risultata positiva al test"}, T^- = \text{"la donna è risultata negativa al test"} \\ G &= \text{"la donna ha una gravidanza"}, G^c = \text{"la donna non ha una gravidanza"} \end{aligned}$$

³⁹ Il testo afferma che "su 1.000 test negativi, 80 erano sbagliati". Questo implica che i restanti 920 test negativi erano esatti, cioè indicavano assenza di gravidanza. Pertanto la probabilità che non vi sia gravidanza, sapendo che il test ha avuto esito negativo, è $\frac{920}{1.000}$. In modo analogo si calcola la probabilità di gravidanza sapendo che l'esito del test è positivo.

Ecco allora come diventa il grafo ad albero:



Ed ecco come si può esprimere in **forma compatta** l'evento "test esatto":

$$(T^- \text{ e } G^c) \text{ o } (T^+ \text{ e } G)$$

Gli errori dei test diagnostici

Il peso che può avere un errore nell'esito del test dipende dalla tipologia del test. Esaminiamo alcuni esempi al riguardo.

- Per un test di gravidanza (effettuato magari con un kit di farmacia) un **Falso Positivo**, può creare un po' d'ansia, ma può venir corretto da una successiva visita medica. Invece un **Falso Negativo** è potenzialmente più grave dato che la madre, in tal caso, non avrebbe motivo di prendere le precauzioni che normalmente si seguono in caso di gravidanza.
- Per un test relativo all'HIV un **Falso Positivo** può risultare molto grave, al punto che si sono verificati casi di suicidio di persone in realtà sane. Un **Falso Negativo** può avere conseguenze ancora più gravi dato che si rassicura il paziente sul suo stato di salute mentre in realtà risulta malato.
- È importante ridurre il numero dei **Falso Positivo** anche negli esperimenti in cui si testa l'efficacia di un farmaco. Altrimenti l'analisi potrebbe sancire erroneamente la bontà di un medicinale che in realtà non ha alcun valore terapeutico.
- Il giudizio di un tribunale è soggetto a valutazioni probabilistiche: è difficile avere la certezza assoluta della colpevolezza o dell'innocenza di un persona. Ma, se da una parte è importante evitare la condanna di un innocente (**Falso Positivo**), dall'altra è altrettanto importante non assolvere un colpevole (**Falso Negativo**).

Ora, se si promuovono misure troppo garantiste per evitare la condanna di innocenti, si rischia di ottenere anche l'assoluzione di qualche colpevole. Il sistema giudiziario italiano ha introdotto un terzo grado di giudizio proprio per tutelare meglio gli innocenti; l'effetto è però anche quello di allungare vistosamente la durata del processo, con il rischio di rilasciare possibili colpevoli per decorrenza dei termini.