



I FONDAMENTALI

*Le parole dell'epidemia spiegate dai
docenti di Scienze Statistiche di Padova*

**Stima della prevalenza di una
malattia con test imperfetto**



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI SCIENZE
STATISTICHE



I FONDAMENTALI

*Le parole dell'epidemia spiegate dai
docenti di Scienze Statistiche di Padova*

**Stima della prevalenza di una
malattia con test imperfetto**



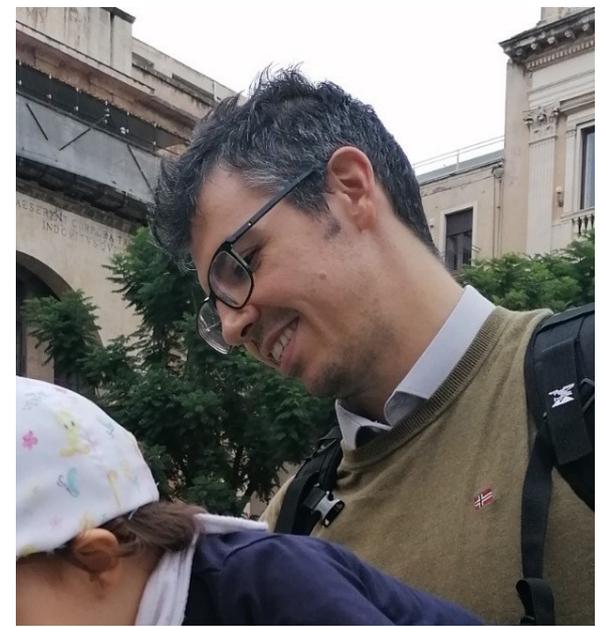
UNIVERSITÀ
DEGLI STUDI
DI PADOVA



CENTRO
DI SCIENZE
STATISTICHE

Prof.ssa
LAURA VENTURA

docente di
STATISTICA MEDICA
e di
MODELLI STATISTICI



Prof.
PAOLO GIRARDI

docente di
MODELLI STATISTICI
APPLICATI



PREVALENZA



La prevalenza $P(M)$ di una malattia è la proporzione di malati presenti in una popolazione.

Una procedura standard per stimare la prevalenza $P(M)$ consiste nel sottoporre a un test diagnostico un campione di n individui rappresentativo della popolazione e di calcolare il rapporto

$$P = \frac{T}{n}$$

dove T è il numero di soggetti risultati positivi al test.

Per la stima della prevalenza da infezione da Coronavirus in Italia, l'Istat ha avviato un'indagine il 25 maggio.

TEST DIAGNOSTICO IMPERFETTO

Il test diagnostico è positivo se segnala la presenza della malattia ed è negativo se non la segnala. Ma in genere il test non produce risultati sicuri al 100%.

Tabella di Errata Classificazione

	Paziente Malato (M)	Paziente Sano (S)
Test Positivo (T ⁺)	Vero Positivo	Falso Positivo
Test Negativo (T ⁻)	Falso Negativo	Vero Negativo

Per la diagnosi del COVID-19 esistono due principali test diagnostici:

- i *tamponi naso/gola* per rilevare l'infezione attiva
 - i *test sierologici* per rilevare l'infezione passata
- In condizioni ideali di laboratorio l'accuratezza di tali test è molto elevata, ma questa si riduce quando vengono applicati.

Sensibilità (*sen*)

$$P(T^+|M)$$

Specificità (*spe*)

$$P(T^-|S)$$

STIMA DELLA PREVALENZA CON TEST IMPERFETTO

Il test è dunque **imperfetto**: sensibilità (*sen*) e specificità (*spe*) non del 100%.

Il test viene applicato su n individui, T dei quali risultano positivi.

La stima $P = T/n$ di $P(M)$ è dunque “distorta”, per via dei Falsi Positivi e dei Falsi Negativi.

Una stima corretta della prevalenza può essere ottenuta a partire dalla tabella di errata classificazione come (*P. Diggle, 2011*):

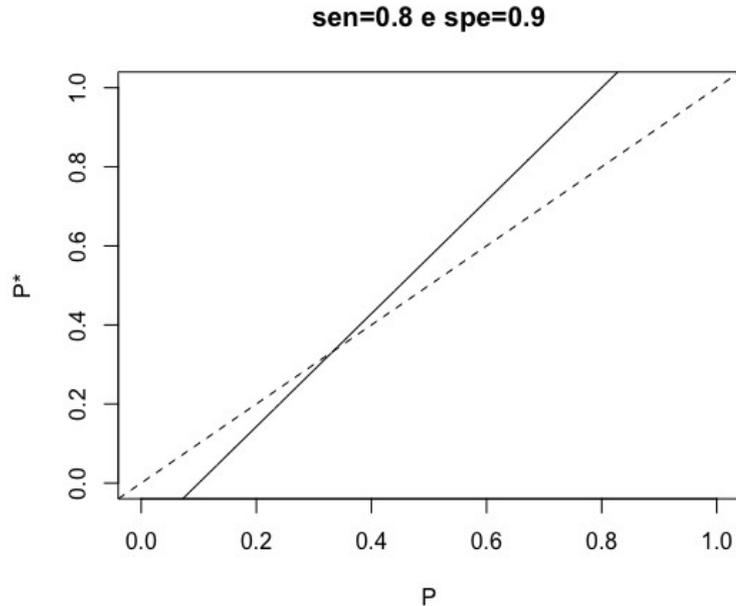
$$P^* = \frac{P - 1 + spe}{(sen + spe - 1)} \quad (\text{per } P > 1 - spe)$$

Quando $sen + spe > 1$, ossia quando il test è preferibile al lancio della moneta, P^* è una funzione crescente di P .

STIMA DELLA PREVALENZA CON TEST IMPERFETTO

Generalmente, per valori bassi della prevalenza si ha $P^* < P$, cioè la stima corretta della prevalenza risulta minore della proporzione calcolata direttamente dai dati osservati.

Ad esempio, nel caso $sen = 0.8$ e $spe = 0.9$ con $P = 0.15$, si ha $P^* = 0.07$, una stima quasi dimezzata del valore osservato.



Nel grafico:

- linea piena: $P^* = (P - 1 + spe) / (spe + sen - 1)$
- linea tratteggiata: $P^* = P$ ($sen = spe = 1$)

Con $sen = 0.8$ e $spe = 0.9$, le linee si intersecano in 0.33.

UN ESEMPIO DI STIMA DELLA PREVALENZA PER IL COVID-19 IN ITALIA

Per il Covid-19, i test diagnostici hanno qualità molto alte in condizioni ideali di laboratorio. Purtroppo quando vengono applicati, la loro accuratezza è inferiore.

Assumiamo $sen = 0.7$ e $spe = 0.99$.
Con i dati della Protezione Civile sui totali positivi e tamponi effettuati, si ha $P = 0.067$.

La stima corretta fornisce

$$P^* = (P - 1 + 0.99)/(0.7 + 0.99 - 1) = 0.084$$

