

Misura kappa di concordanza, introdotta da Cohen

Prof. Giuseppe Verlato

Sezione di Epidemiologia e Statistica Medica
Università degli Studi di Verona

Esiste un *gold standard* (riferimento aureo)?

SI'

Si calcola la **SENSIBILITÀ**
e la **SPECIFICITÀ** del test
contro il gold standard

NO: si calcola la misura
KAPPA di CONCORDANZA,
introdotta da **COHEN**

Esempio:

Nella diagnosi del cancro colo-
rettale il gold standard è la
colonscopia + biopsia

*Inter-rater o intra-rater
agreement:* concordanza fra i
giudizi espressi da operatori
diversi o dallo stesso operatore

Studio sulla discinesia tardiva

		giudice b		<i>totale</i>
		<i>presente</i>	<i>assente</i>	
giudice a	<i>pres.</i>	123	10	133
	<i>ass.</i>	6	29	35
	<i>totale</i>	129	39	168

$$\chi^2 = 88.23 \longrightarrow P < 0.0000001$$

Bergen et al (1992) Studio sulla discinesia tardiva.

Il chi-quadrato ci dà scarse informazioni.

Abbiamo semplicemente falsificato l'ipotesi nulla che le diagnosi del primo clinico fossero del tutto indipendenti dalle diagnosi del secondo clinico.

Tuttavia questo risultato può essere conseguito anche da due clinici molto
MEDIOCRI.

Misura di concordanza osservata (Proportion of agreement)

		giudice b		totale
		presente	assente	
giudice a	pres.	a	b	a+b
	ass.	c	d	c+d
totale		a+c	b+d	N

$$P_0 = \frac{a + d}{N}$$

Misura di concordanza osservata (Proportion of agreement)

		giudice b		totale
		presente	assente	
giudice a	pres.	123	10	133
	ass.	6	29	35
totale		129	39	168

$$P_0 = \frac{123 + 29}{168} = \frac{152}{168} = 0.905$$

Bergen et al (1992) Studio sulla discinesia tardiva.

Però anche due persone **completamente ignoranti** di medicina riescono a raggiungere una certa concordanza, per semplice effetto del caso.

A noi interessa valutare di quanto la concordanza fra gli **esperti** eccede la concordanza attesa per semplice effetto del caso.

**Concordanza ATTESA
per effetto del caso
(agreement expected just by chance)**

$$P_{\text{exp}} = \frac{a_{\text{exp}} + d_{\text{exp}}}{N}$$

Sotto l'ipotesi di indipendenza statistica

atteso = (totale riga) * (totale colonna) / (totale generale)

		giudice b		
		<i>presente</i>	<i>assente</i>	<i>totale</i>
giudice a	<i>pres.</i>	a	b	a+b
	<i>ass.</i>	c	d	c+d
<i>totale</i>		a+c	b+d	N

$$a_{\text{exp}} = \frac{(a+b) * (a+c)}{N}$$

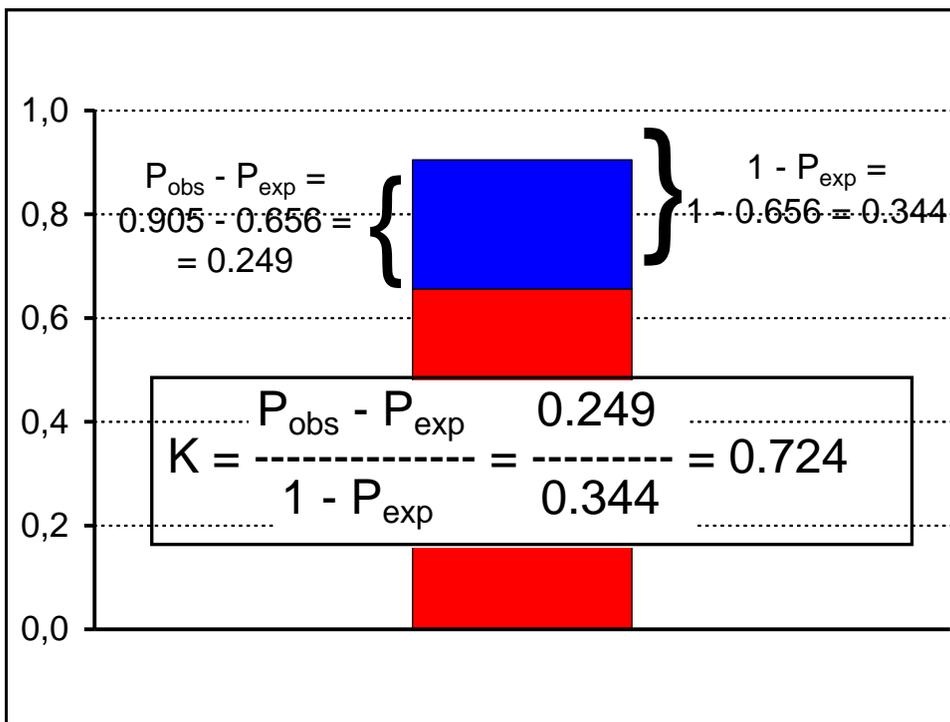
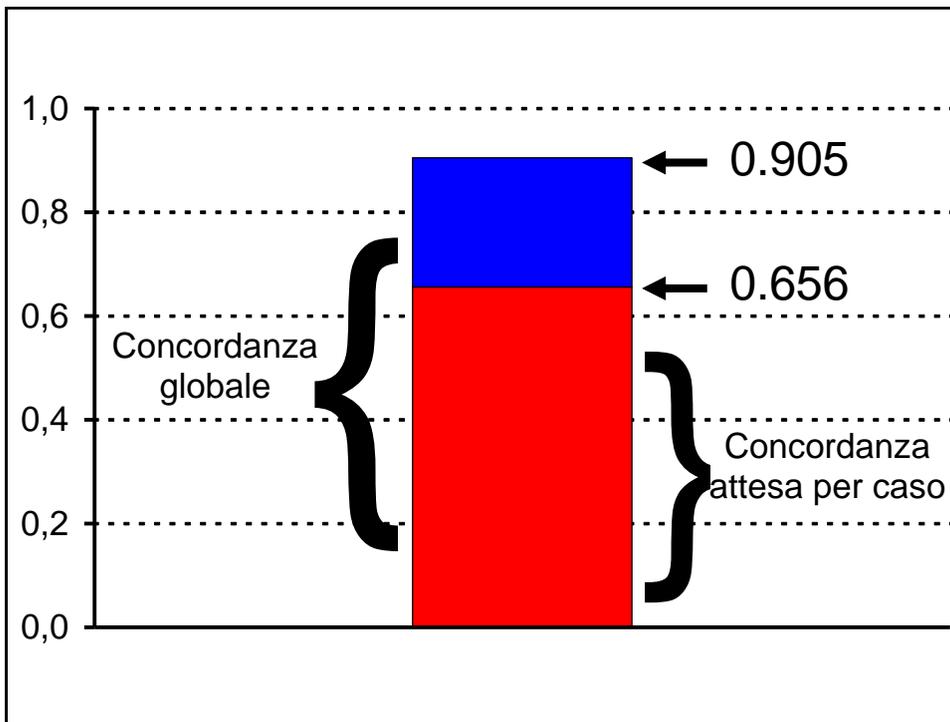
Concordanza attesa

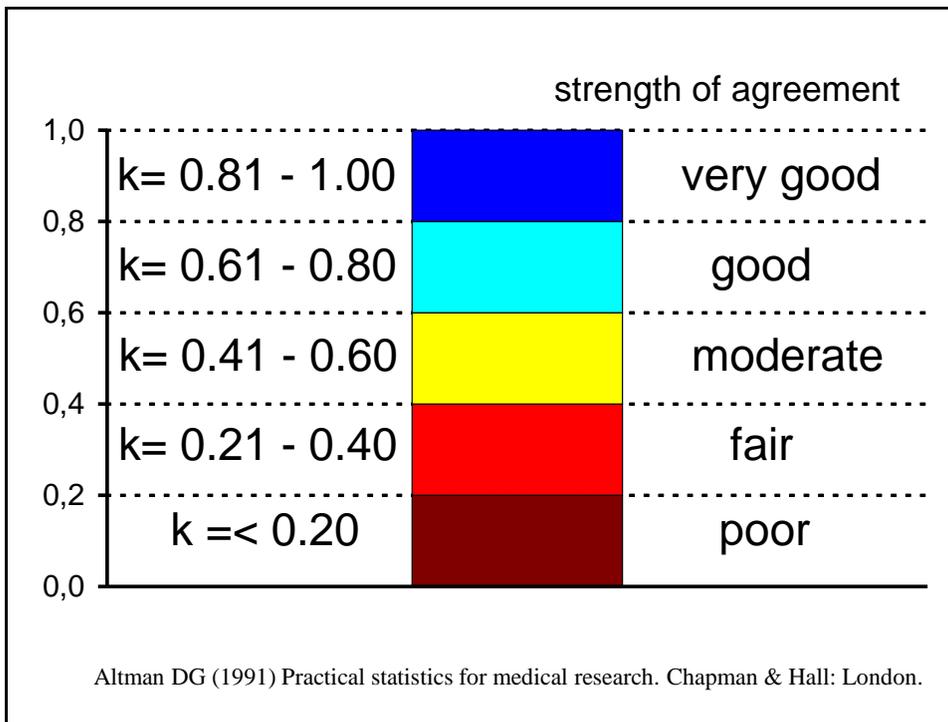
		giudice b		
		<i>presente</i>	<i>assente</i>	<i>totale</i>
giudice a	<i>pres.</i>	123	10	133
	<i>ass.</i>	6	29	35
<i>totale</i>		129	39	168

$$a_{\text{atteso}} = (133 * 129) / 168 = 102.125$$

$$d_{\text{atteso}} = (35 * 39) / 168 = 8.125$$

$$P_e = \frac{102.1 + 8.1}{168} = \frac{110.25}{168} = 0.656$$





A parità di concordanza, la misura k è massima quando la prevalenza dei positivi (malati) è prossima a 0.50.

		giudice b			
		presente	assente	totale	
giudice a	pres.	70	10	80	$P_{obs} = 80 / 100 = 0.80$ $a_{exp} = (80 \cdot 80) / 100 = 64$ $d_{exp} = (20 \cdot 20) / 100 = 4$ $P_{exp} = (64 + 4) / 100 = 0.68$ $k = (0.80 - 0.68) / (1 - 0.68) = 0.38$
	ass.	10	10	20	
	totale	80	20	100	
		giudice b			
		presente	assente	totale	
giudice a	pres.	40	10	50	$P_{obs} = 80 / 100 = 0.80$ $a_{exp} = (50 \cdot 50) / 100 = 25$ $d_{exp} = (50 \cdot 50) / 100 = 25$ $P_{exp} = (25 + 25) / 100 = 0.50$ $k = (0.80 - 0.50) / (1 - 0.50) = 0.60$
	ass.	10	40	50	
	totale	50	50	100	

E' possibile studiare separatamente la concordanza fra le risposte positive e le risposte negative, attraverso gli indici:

P_{pos} = proportion of positive agreement

P_{neg} = proportion of negative agreement

		2 nd survey			
		No	Yes	Total	
First survey	No	a	b	a+b	$P_{\text{neg}} = \frac{\mathbf{a}}{(\mathbf{a+b+a+c})/2}$
	Yes	c	d	c+d	
	Total	a+c	b+d	N	

		2 nd survey			
		No	Yes	Total	
First survey	No	a	b	a+b	$P_{\text{pos}} = \frac{\mathbf{d}}{(\mathbf{b+d+c+d})/2}$
	Yes	c	d	c+d	
	Total	a+c	b+d	N	

Cicchetti & Feinstein, J Clin Epidemiol, 1990