Probabilità (definizione classica)

Se un evento si può verificare in *N* modi mutuamente esclusivi ed ugualmente possibili, e se *m* di questi possiede la caratteristica *E*, la probabilità di verificarsi di *E* è:

$$P(E) = \frac{m}{N}$$

Esempio

- Se si lancia un dado a 6 facce non truccato, la probabilità che si osservi 2 è 1/6 ed è uguale per le rimanenti facce.
- Se si sceglie a caso una carta da un mazzo di 52 carte ben mescolato, la probabilità di estrarre una carta di cuori è 13/52.

Probabilità (definizione frequentista)

Se si ripete un esperimento un gran numero di volte *n* e se un certo evento con caratteristica *E* si verifica *m* volte, la frequenza relativa di successo di *E*, *m/n*, sarà approssimativamente uguale alla probabilità di *E*:

$$P(E) \approx \frac{m}{n}$$

Esempio

- Supponiamo di effettuare n=100 lanci di una moneta non truccata e di osservare m=47 volte Testa. Allora la probabilità di osservare Testa è approssimativamente 47/100.

Definizioni

- Si dice *evento certo* o *spazio degli eventi,* e si denota con *S,* l'insieme di tutti i risultati di un esperimento.
- Gli elementi di *S* sono chiamati *risultati sperimentali*.
- Si dice *evento E* un qualsiasi sottoinsieme di *S* e si denota con $E \subset S$.
- Un evento si dice elementare quando è costituito da un solo elemento di *S*.
- L'evento impossibile $\emptyset \subset S$.

Esempio

Indichiamo con f_i le 6 facce di un dado: $S=\{f_1, f_2, \ldots, f_6\}$.

Un risultato sperimentale è f_2.

L'evento $\{f_2\}$ è un evento elementare.

$\{f_1, f_2, f_3\}$ è un evento.

$\emptyset, \{f_1\}, \ldots, \{f_1, f_2\}, \ldots, \{f_1, f_2, f_3\}, \ldots, S$ sono tutti i possibili sottoinsiemi di S.

S ha $2^6=64$ sottoinsiemi.

Richiami di insiemistica

Siano A e B sottoinsiemi di S.

$A \cup B$ A **unione** B è l'insieme degli elementi di S che appartengono ad A oppure a B.

$A \cap B$ A **intersezione** B è l'insieme degli elementi di S che appartengono sia ad A che a B.

Se $A \cap B = \emptyset$ gli insiemi si dicono **disgiunti**

in probabilità si dicono **mutuamente esclusivi** o **incompatibili**

$\overline{A}$ **complementare** di A è l'insieme degli elementi di S che non appartengono ad A.

Il concetto di prova

Una *prova* è una singola esecuzione di un esperimento. Ad ogni prova osserviamo un singolo risultato f_i.

Diremo che un evento A *si verifica* durante questa prova se esso contiene f_i. S si verifica ad ogni prova e Ø non si verifica mai.

$A \cup B$ si verifica se A o B o entrambi si verificano.

$A \cap B$ si verifica se sia A che B si verificano.

Se A e B sono mutuamente e A si verifica, allora B non si verifica.

Se $A \subset B$ e A si verifica, allora B si verifica.

Ad ogni prova A oppure $\overline{A}$ si verifica.

Esempio

Il nostro esperimento consiste nel lancio di un dado. In una prova, ossia in un lancio di un dado, osserviamo il risultato f_5. Allora:

1) l'evento $\{f_5\}$ si verifica;

2) l'evento $\{dispari\}$ si verifica.

Un totale di 32 eventi si verificano.

Assiomi

Ad ogni evento *A* assegniamo un numero *P(A)* che chiamiamo la probabilità di *A* che soddisfa le seguenti condizioni:

1) $P(A) \geq 0$.
2) $P(S) = 1$.
3) Se $A \cap B = \emptyset$ allora $P(A \cup B) = P(A) + P(B)$.

Proprietà 1

La probabilità dell'evento impossibile è zero:

$$P(\varnothing) = 0.$$

Dimostrazione. Infatti, per ogni evento A, si ha:

$$A \cap \varnothing = \varnothing \text{ e } A \cup \varnothing = A.$$

Allora per l'assioma 3 si ha:

$P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing).$

Quindi $P(\varnothing) = 0.$

Proprietà 2

Per ogni evento A risulta:

$$P(A) = 1 - P(\bar{A}) \quad \text{e} \quad P(A) \leq 1.$$

Dimostrazione. Infatti,

$A \cup \bar{A} = S \quad$ e $\quad A \cap \bar{A} = \emptyset$. Allora:

$1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

Quindi:

$P(A) = 1 - P(\bar{A})$. Inoltre essendo per l'assioma 1 $P(\bar{A}) \geq 0$ segue che $P(A) \leq 1$.

Proprietà 3

Per ogni evento A e B risulta:

1) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;

2) $P(A \cup B) \leq P(A) + P(B)$.

Dimostrazione. Infatti, gli eventi $A \cup B$ e B si possono scrivere come unione di due eventi mutuamente esclusivi:

$A \cup B = A \cup (\bar{A} \cap B) \Rightarrow P(A \cup B) = P(A) + P(\bar{A} \cap B)$.

$B = (A \cap B) \cup (\bar{A} \cap B) \Rightarrow P(B) = P(A \cap B) + P(\bar{A} \cap B)$.

Dalla seconda segue $P(\bar{A} \cap B) = P(B) - P(A \cap B)$. Sostituendo questa nella prima discende la 1).

Proprietà 4

Gli eventi A e B sono tali che $B \subset A$. Allora:

1) $P(A) = P(B) + P(A \cap \overline{B})$;

2) $P(B) \leq P(A)$.

Dimostrazione. Infatti,

$A = B \cup (A \cap \overline{B})$. Inoltre, $B \cap (A \cap \overline{B}) = \emptyset$.

Allora per l'assioma 3 risulta:

$P(A) = P(B) + P(A \cap \overline{B})$.

Spazio degli eventi finito

Sia S costituito da un numero finito N di risultati sperimentali:

$$S = \{f_1, f_2, \ldots, f_N\}.$$

Indichiamo con p_i la probabilità di ogni singolo evento elementare $\{f_i\}$:

$$P\{f_i\} = p_i.$$

Per gli assiomi della probabilità si ha:

$$p_i \geq 0 \qquad p_1 + p_2 + \ldots + p_N = 1.$$

Spazio degli eventi finito (cont.)

Sia A un evento costituito da k elementi:

$$A = \{f_1, f_2, \ldots, f_k\}.$$

In questo caso A può essere scritto come l'unione di eventi elementari $\{f_i\}$:

$$A = \{f_1\} \cup \{f_2\} \cup \ldots \cup \{f_k\}.$$

Quindi:

$$P(A) = P\{f_1\} + P\{f_2\} + \ldots + P\{f_k\}$$
$$= p_1 + p_2 + \ldots + p_k.$$

Caso particolare

Nell'ipotesi in cui le probabilità p_i degli eventi elementari sono tutte uguali, allora:

$$p_i = \frac{1}{N}$$

e

$$P(A) = \frac{k}{N}$$

Esempio 1

Ricerca sull'uso della cocaina da parte degli uomini e delle donne (Erickson & Murray '89).

Num. volte in cui è stata usata la cocaina	**Maschio (M)**	**Femmina (F)**	**Totale**
1-19 volte (A)	**32**	**7**	**39**
20-99 volte (B)	**18**	**20**	**38**
100 volte (C)	**25**	**9**	**34**
Totale	**75**	**36**	**111**

Esempio 1 (cont.)

Nota: gli autori affermano che il campione è rappresentativo di utilizzatori tipici: i soggetti non sono né in trattamento, né in prigione.

Qual è la probabilità che la persona scelta sia un maschio?

Soluzione: assumiamo che M e F siano categorie mutuamente esclusive e che ogni persona abbia la stessa probabilità di essere scelta. Allora la probabilità di essere M è il numero di soggetti con la caratteristica M diviso il numero totale di soggetti:

$P(M)$ = Numero di maschi / Numero totale di soggetti

= 75 / 111 = 0.6757

$P(M)$ prende il nome di *probabilità marginale*.

Esempio 2

Qual è la probabilità che una persona scelta a caso dai 111 soggetti sia maschio (M)

e

abbia usato la cocaina un numero di volte maggiore o uguale a 100 (C) nella sua vita?

Soluzione: l'evento che ci interessa è $M \cap C$ che si verifica quando gli eventi M e C si verificano *congiuntamente*. $M \cap C$ è costituito da tutti i soggetti maschi che hanno usato cocaina un numero di volte maggiore o uguale a 100. Quindi:

$$P(M \cap C) = 25 / 111 = 0.2252.$$

$P(M \cap C)$ prende il nome di *probabilità congiunta*.

Probabilità condizionata

Consideriamo due eventi *A* e *B* e supponiamo di sapere che l'evento *B* si è verificato. Il verificarsi di *B* modifica la probabilità di *A*?

La probabilità di un evento *A* dato che l'evento *B* si è verificato prende il nome di *probabilità condizionata* e si denota con:

$$P(A \mid B).$$

Esempio 3

Supponiamo di scegliere a caso un soggetto tra i 111 soggetti e che esso sia maschio (M). Qual è la *P* che questo maschio abbia usato cocaina un numero di volte maggiore o uguale a 100 durante la sua vita (C) ?

Soluzione: avendo estratto un maschio, siamo interessati al consumo di cocaina trai maschi e non tra tutti i 111 soggetti. Le femmine sono escluse. Quindi l'evento $C \mid M$ è l'insieme dei maschi che ha usato cocaina un numero di volte maggiore o uguale a 100. Quindi:

$$P(C \mid M) = 25 / 75 = 0.33.$$

Domanda

Quale relazione intercorre tra la probabilità congiunta di due eventi A e B

$$P(A \cap B)$$

e la probabilità condizionata di *A* dato *B*

$$P(A \mid B)?$$

Definizione

La probabilità condizionata di A dato B è per definizione il rapporto:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Assumendo che *P(B)* sia diversa da zero.

Osservazione

La probabilità condizionata ci consente di calcolare in modi diversi la probabilità congiunta. Siano A e B due eventi. Allora:

$$P(A \cap B) = P(A \mid B)\, P(B).$$

Oppure:

$$P(A \cap B) = P(B \mid A)\, P(A).$$

Interpretazione

In n ripetizioni di un esperimento, indichiamo con n_A, n_B e $n_{A \cap B}$ il numero di occorrenze degli eventi A, B e $A \cap B$ rispettivamente. Allora:

$$P(A) \approx \frac{n_A}{n} \qquad P(B) \approx \frac{n_B}{n} \qquad P(A \cap B) \approx \frac{n_{A \cap B}}{n}$$

Quindi:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \approx \frac{n_{A \cap B} / n}{n_B / n} = \frac{n_{A \cap B}}{n_B}$$

Interpretazione (cont.)

Se scartiamo tutte le prove nelle quali l’evento *B* non si verifica e prendiamo solo la sottosequenza di prove in cui l’evento *B* si verifica, allora *P(A | B)* è uguale alla frequenza relativa dell’occorrenza dell’evento *A* in quella sottosequenza, ossia

$$n_{A \cap B}/n_B.$$

Semplici proprietà

1) Se $A \subset B$ allora $P(B \mid A) = 1$.

Infatti, se $A \subset B$ allora $A \cap B = A$. Quindi:

$P(B \mid A) = P(B \cap A) / P(A) = P(A) / P(A) = 1$.

2) Se $A \subset B$ allora $P(A \mid B) \geq P(A)$.

Infatti, per la proprietà 2 si ha: $P(B) \leq 1$.

Inoltre, se $A \subset B$ allora $A \cap B = A$. Quindi:

$$P(A \mid B) = P(A \cap B) / P(B)$$

$$= P(A) / P(B) \geq P(A).$$

Esempio 4

Risolvere l'esempio 3 utilizzando la definizione di probabilità condizionata.

Soluzione:

$$P(C \mid M) = P(C \cap M) \,/\, P(M)$$

$$= (25/111) \,/\, (75/111)$$

$$= 25 \,/\, 75 = 0.33.$$

Esempio 5

Scegliamo a caso una persona tra i 111 soggetti dell'esempio 1. Qual è la probabilità che questa persona si maschio (M) oppure femmina (F)?

Soluzione: poiché i due generi sono mutuamente esclusivi allora: $P(M \cup F) = P(M) + P(F) = 75/111 + 36/111 = 1$.

Esempio 6

Scegliamo un soggetto tra i 111 dell'esempio 1. Qual è la probabilità che questo soggetto sia maschio (M) oppure abbia usato cocaina un numero di volte maggiore o uguale a 100 nella sua vita (C)?

Soluzione: in questo caso gli eventi *M* e *C* non sono mutuamente esclusivi e quindi:

$$P(M \cup C) = P(M) + P(C) - P(M \cap C)$$
$$= 75/111 + 34/111 - 25/111$$
$$= 84/111 = 0.7568.$$

Eventi indipendenti

Due eventi A e B si dicono indipendenti se:

$$P(A \cap B) = P(A)\, P(B).$$

Come conseguenza si ha:

$$P(A \mid B) = P(A).$$

E inoltre:

$$P(B \mid A) = P(B).$$

Esempio 7

In un campione di 100 studenti (60 *F* e 40 *M*) 40 portano gli occhiali (24 O_F e 16 O_M). La probabilità *P(O)* che uno studente porti gli occhiali è: $P(O) = 40/100 = 0.4$

Calcoliamo:

$$P(O \mid M) = P(O \cap M) / P(M)$$

$$= (16/100) / (40/100) = 0.4$$

Quindi l'informazione aggiuntiva che lo studente è un ragazzo non muta la probabilità che uno studente porti gli occhiali.

Osservazione

Se A e B sono *eventi indipendenti* allora:

$$P(A \mid B) = P(A).$$

Se A e B sono *eventi mutuamente* esclusivi allora:

$$P(A \mid B) = 0$$

in quanto $P(A \cap B) = 0$.

Teorema delle probabilità totali

Osservazione: in generale un evento A può essere espresso come unione di eventi mutuamente esclusivi.

Siano $B_1, B_2, \dots, B_n$ n eventi tali che:

$$B_i \cap B_j = \emptyset \text{ per } i \neq j \qquad \text{e} \qquad B_1 \cup B_2 \cup \dots \cup B_n = S.$$

L'evento A si può scrivere come:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n).$$

Poiché gli eventi $A \cap B_i$ sono mutuamente esclusivi, allora per l'assioma 3 si ha:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_n).$$

Sapendo che $P(A \cap B_i) = P(A|B_i)\, P(B_i)$ allora:

$$P(A) = P(A|B_1)\, P(B_1) + P(A|B_2)\, P(B_2) + \dots + P(A|B_n)\, P(B_n).$$

Questo risultato è noto con il nome di teorema delle probabilità totali.

Esempio 8

Nel contesto dell'esempio 1, consideriamo l'evento maschio *(M)*. Tale evento si può verificare congiuntamente con ciascuno dei tre eventi *(A)*, *(B)* e *(C)*. Poiché questi ultimi sono mutuamente esclusivi e $A \cup B \cup C = S$ allora:

$$M = (M \cap A) \cup (M \cap B) \cup (M \cap C).$$

Per l'assioma 3 segue:

Esempio 8 (cont.)

$$P(M) = P(M \cap A) + P(M \cap B) + P(M \cap C).$$

$$P(M) = 32/111+18/111+25/111=75/111.$$

Per il teorema delle probabilità totali si ha:

$P(M) = P(M \mid A)\ P(A) + P(M \mid B)\ P(B) + P(M \mid C)\ P(C).$

$P(M) = (32/39)\ (39/111) + (18/38)\ (38/111) +$
$\qquad (25/34)\ (34/111).$

$P(M) = 32/111+18/111+25/111=75/111.$

Teorema di Bayes

Nota: siano *A* e *B* due eventi. E' noto che:

$$P(A \cap B) = P(A \mid B)\, P(B)$$

e

$$P(A \cap B) = P(B \mid A)\, P(A).$$

Quindi:

$$P(B \mid A)\, P(A) = P(A \mid B)\, P(B),$$

da cui:

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)}$$

Teorema di Bayes (cont.)

Siano $B_1, B_2, \ldots, B_n$ n eventi tali che:

$$B_i \cap B_j = \emptyset \text{ per } i \neq j \qquad \text{e} \qquad B_1 \cup B_2 \cup \ldots \cup B_n = S.$$

Allora per la nota precedente risulta che per $i=1,2,\ldots,n$:

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{P(A)}$$

Per il teorema delle probabilità totali si ha:

$$P(A) = P(A|B_1)\, P(B_1) + P(A|B_2)\, P(B_2) + \ldots + P(A|B_n)\, P(B_n)$$

da cui:

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{P(A|B_1)\, P(B_1) + P(A|B_2)\, P(B_2) + \ldots + P(A|B_n)\, P(B_n)}$$

Esempio 9

La National Health Interview Survey condusse un'indagine su casi di alterazione delle capacità uditive dovuti ad infortuni in soggetti di età ≥ 17 anni negli anni 1980-1981.

Posizione lavorativa	Soggetti con alterazioni capacità uditive (C)	Soggetti sani (H)	Totale
Impiegati *(E_1)*	552	98365	98917
Disoccupati *(E_2)*	27	7435	7462
Non nella forza lavoro *(E_3)*	368	56410	56778
Totale	947	162210	163157

Esempio 9 (cont.)

Calcoliamo la probabilità che un soggetto sia impiegato *(E_1)* dato che lo stesso sia affetto da un'alterazione dell'udito *(C)*: $P(E_1 \mid C)$.

Applicando il teorema di Bayes si ha:

$$P(E_1 \mid C) = \frac{P(C \mid E_1)\, P(E_1)}{P(C)}$$

$$P(E_1 \mid C) = \frac{(552/98917)\ (98917/163157)}{(947/163157)}$$

$$P(E_1 \mid C) = 552/947 = 0.583.$$

Questo risultato si può ricavare direttamente dalla tabella.

Esempio 9 (cont.)

Nota che l'informazione supplementare che un determinato soggetto abbia alterazioni uditive *(C)* cambia la probabilità che questo sia un impiegato (E_1). Infatti:

$$P(E_1 \mid C) = 0.583$$

mentre

$$P(E_1) = 98917/163157 = 0.606.$$

Aver saputo che un soggetto è un caso riduce la probabilità che esso sia impiegato.

Test diagnostici

Un *test diagnostico* o di *screening* è un procedimento che si applica in genere a soggetti sani per determinare la presenza in essi di una patologia o per stabilire una loro eventuale suscettibilità (predisposizione) alla patologia.

Indichiamo con H (healthy) l'evento che un soggetto è sano e con D (diseased) l'evento che un soggetto è malato. Allora:

$$H \cap D = \emptyset \quad \text{e} \quad H \cup D = S.$$

Inoltre, indichiamo con P l'evento che un soggetto è risultato positivo al test e con N l'evento che un soggetto è risultato negativo al test. Anche in questo caso:

$$P \cap N = \emptyset \quad \text{e} \quad P \cup N = S.$$

Vogliamo determinare:

$$P(D \mid P) \quad \text{e} \quad P(H \mid N)$$

Esito del test

Stato del soggetto		P	N
	D	TP	FN
	H	FP	TN

TP (True Positive): soggetti positivi al test che sono malati.

FN (False Negative): soggetti negativi al test che sono malati.

FP (False Positive): soggetti positivi al test che sono sani.

TN (True Negative): soggetti negativi al test che sono sani.

Indicatori di un test diagnostico

Si definisce *sensibilità* di un test diagnostico:

$$Se = P(P \mid D).$$

Quindi:

$$Se = \frac{TP}{TP + FN}$$

Si definisce *specificità* di un test diagnostico:

$$Sp = P(N \mid H).$$

Quindi:

$$Sp = \frac{TN}{TN + FP}$$

Esempio 10

La probabilità di recupero in soggetti affetti da cancro della cervice uterina è elevata in caso di individuazione precoce.

Il Pap test è una procedura di screening che può individuare il cancro anche in casi asintomatici.

Studi condotti negli anni 1972-1978 in 306 laboratori in 44 stati hanno rilevato che:

$P(N \mid D)$ = 0.1625 (probabilità di un FN);

$P(P \mid D)$ = 0.8375 (*Se*).

$P(P \mid H)$ = 0.1864 (probabilità di un FP);

$P(N \mid H)$ = 0.8136 (*Sp*).

Osservazioni

Questi sono indicatori della bontà di un test diagnostico. Infatti per la loro stima il test si applica a soggetti il cui fenotipo è noto *a priori*.

Un test con elevata *Se* è preferibile quando non si vuole perdere nessun malato ($FN \approx 0$).

Un test con elevata *Sp* è preferibile quando può essere dannoso avere falsi positivi.

Determiniamo $P(D \mid P)$

Applichiamo il teorema di Bayes:

$$P(D \mid P) = \frac{P(P \mid D)\,P(D)}{P(P \mid D)\,P(D) + P(P \mid H)\,P(H)}$$

Sapendo che le donne affette da cancro della cervice sono 8.3 su 100.000 negli anni ’82-’83:

$P(D) = 0.000083$ e $P(H) = 1 - P(D) = 0.999917$.

Allora: $P(D \mid P) = 0.000373$.

$P(D \mid P)$ è detto *valore predittivo di un test positivo*.

Determiniamo $P(H \mid N)$

Applichiamo il teorema di Bayes:

$$P(H \mid N) = \frac{P(N \mid H)\, P(H)}{P(N \mid D)\, P(D) + P(N \mid H)\, P(H)}$$

Nelle stesse ipotesi precedenti si ha:

$P(H \mid N) = 0.999983$.

$P(H \mid N)$ è detto *valore predittivo di un test negativo*.

Questo indica che su 1.000.000 di donne risultate negative al Pap test, 999983 erano sane.

Rischio relativo

Il rischio relativo (*RR*) è il rapporto tra la probabilità di sviluppare la patologia in soggetti esposti e la probabilità di sviluppare la patologia in soggetti non esposti:

$$RR = \frac{P(D \mid esposto)}{P(D \mid non\ esposto)}$$

Nota che con il termine esposto si vuole indicare una qualsiasi feature che caratterizza un gruppo di soggetti (esempio di mutazioni in alleli di pazienti affetti da CD).

Esempio 11

In uno studio di coorte per analizzare i fattori di rischio del cancro alla mammella, una donna è considerata “esposta” se ha partorito il primo bambino ad una età ≥ 25 anni.

	D	HC	
Età ≥ 25 anni	31	1597	1628
Età < 25 anni	65	4475	4540
Totale	96	6072	6168

Esempio 11 (cont.)

$P(D \mid esposto) = 31 / 1628.$

$P(D \mid non\ esposto) = 65 / 4540.$

Allora:

$$RR = \frac{31 / 1628}{65 / 4540}$$

$$RR = 1.33$$

Quindi le donne esposte hanno una probabilità 1.33 volte maggiore di sviluppare un cancro alla mammella di quelle non esposte.

Osservazione

1) $RR = 1$ implica che

$$P(D \mid esposto) = P(D \mid non\ esposto)$$

ossia non c'è associazione tra malattia ed esposizione.

2) $RR > 1$ implica che

$$P(D \mid esposto) > P(D \mid non\ esposto)$$

ossia che l'esposizione ha un ruolo di suscettibilità rispetto alla patologia.

3) $RR < 1$ implica che

$$P(D \mid esposto) < P(D \mid non\ esposto)$$

ossia che l'esposizione ha un ruolo protettivo rispetto alla patologia.

Rischio ed eventi rari

Uno studio condotto negli Stati Uniti su uomini di età ≥ 35 anni ha mostrato che:

P(morte per cancro al polmone | fumatore) = 0.002679

P(morte per cancro al polmone | non fumatore) = 0.000154

con

$$RR = 17.4$$

Quindi anche se la probabilità dell' evento considerato è bassa (evento raro), il rischio relativo mette in evidenza l'effetto del fumo sulla probabilità che un soggetto muoia di cancro al polmone.

Odds

Se un evento *A* si verifica con probabilità *p*, l'odds (pronostico) in favore dell'evento A è $p / (1 - p)$ a 1.

Nota: l'odds di un evento è il rapporto tra la probabilità che esso si verifichi e la probabilità che esso non si verifichi.

Se p = 2/3 allora l'odds è (2/3)/(1-2/3)=2 quindi la probabilità che l'evento si verifichi è due volte quella che non si verifichi.

Odds Ratio

L'odds ratio (OR) è il rapporto tra l'odds della malattia tra i soggetti esposti e l'odds della malattia tra i soggetti non esposti:

$$OR = \frac{P(D \mid esposto) \,/\, [1 - P(D \mid esposto)]}{P(D \mid non\ esposto) \,/\, [1 - P(D \mid non\ esposto)]}$$

o in modo equivalente:

$$OR = \frac{P(esposizione \mid D) \,/\, [1 - P(esposizione \mid D)]}{P(esposizione \mid H) \,/\, [1 - P(esposizione \mid H)]}$$

Esempio 12

Studio caso-controllo per analizzare l'influenza dell'uso di contraccettivi orali (CO) nel cancro alla mammella.

	D	HC	
Uso CO	273	2641	2914
Non uso CO	716	7260	7976
Totale	989	9901	10890

$$OR = \frac{(273/989) \,/\, [1 - 273/989]}{(2641/9901) \,/\, [1 - 2641/9901]} = 1.05$$

L'uso di CO non è associato al cancro alla mammella.