LEZIONI DI STATISTICA MEDICA

Teoria della probabilità

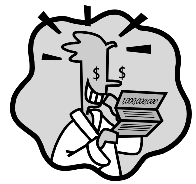
Sezione di Epidemiologia & Statistica Medica
Università degli Studi di Verona

ELEMENTI DI TEORIA DELLA PROBABILITA'

La TEORIA DELLA PROBABILITA' ci permette di studiare e descrivere i **fenomeni aleatori**.

***DEFINIZIONE*:** un fenomeno è aleatorio quando di esso non si può predire con certezza il risultato.

Esempi:
- *Il risultato dell'estrazione al lotto*
- *Il risultato del lancio di una moneta*
- *Il risultato dell'esposiz. al bacillo di Cock*
- *Il risultato dell'esposiz. al fumo di sigaretta*



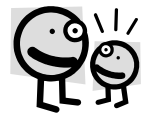
EVENTI e SPAZI CAMPIONARI

Un **esperimento** è un qualsiasi processo di osservazione o misurazione.

Esempi:
- *estrazione di un numero al lotto*
- *lancio di una moneta*
- *valutazione della presenza di un'infezione virale*
- *misura dell'altezza su un campione di bambini*

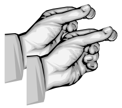
Ad ogni esperimento è associato uno **spazio campionario S** costituito dall'insieme dei possibili risultati. I singoli risultati dell'esperimento sono detti **elementi di S** o **eventi elementari**.

Per la rappresentazione degli spazi campionari e dei loro elementi si utilizza la NOTAZIONE INSIEMISTICA.



ESPERIMENTO	SPAZIO CAMPIONARIO
lancio di un dado	{1, 2, 3, 4, 5, 6}
misurazione della temp. corporea	$\{x \mid 34< x <42\}$
lancio consecutivo di due monete	{TT, TC, CT, CC}
misurazione del sesso e del fumo in un passante	{Mfum, M nn fum, F fum, F nn fum}

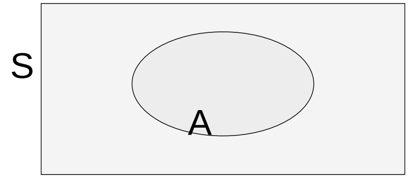
E' possibile definire un qualsiasi evento A come combinazione di più eventi elementari
$\rightarrow$ **EVENTO COMPOSTO**

SPAZIO CAMPIONARIO (S)		**EVENTO (A)**	
	{TT, TC, CT, CC}	*'almeno una testa nel lancio di due monete'*	= {TT, TC, CT}
	{1, 2, 3, 4, 5, 6}	*'numero dispari nel lancio di un dado'*	= {1, 3, 5}

Ogni evento A e' un **SOTTOINSIEME** di S:

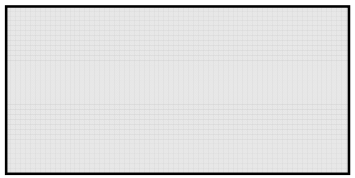
$$A \subset S$$

Gli eventi e gli spazi campionari possono essere rappresentati graficamente mediante i DIAGRAMMI DI VENN:

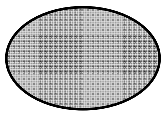


dove l'evento A è il sottoinsieme formato dagli eventi elementari in esso inclusi.

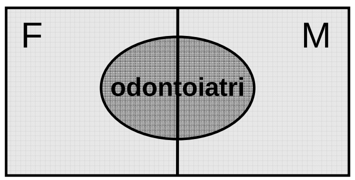
DIAGRAMMI DI VENN



Intera popolazione in studio
(studenti della facoltà di medicina)



sottogruppo (odontoiatri)



N.B.

Un evento è CERTO se comprende tutti gli elementi di S → $A = S$

Un evento è IMPOSSIBILE se non comprende alcun elemento di S → $A = \emptyset$

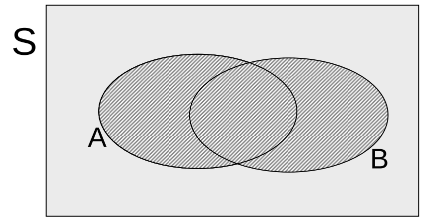
OPERAZIONI LOGICHE SUGLI EVENTI

Dato uno spazio campionario S e degli eventi A_i in esso inclusi, è possibile **definire** mediante operazioni logiche **nuovi eventi**.

Il valore di questi nuovi eventi rimane determinato noto quello degli eventi dati.

UNIONE DI EVENTI - SOMMA LOGICA: $A \cup B$

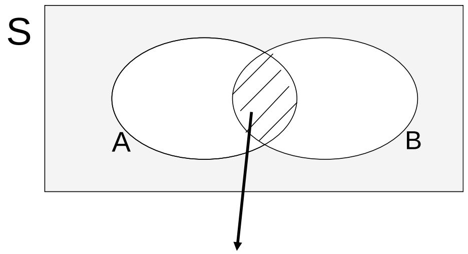
Siano A e B due eventi associati ad un esperimento: l'evento C è definito **unione di A e B** se comprende tutti gli elementi che appartengono ad A **<u>oppure</u>** a B.



$C = A \cup B$

INTERSEZIONE DI EVENTI - PRODOTTO LOGICO: $A \cap B$

Siano A e B due eventi associati ad un esperimento: l'evento C è definito **intersezione di A e B** se comprende tutti gli elementi che appartengono ad A **<u>e contemporaneamente</u>** a B.



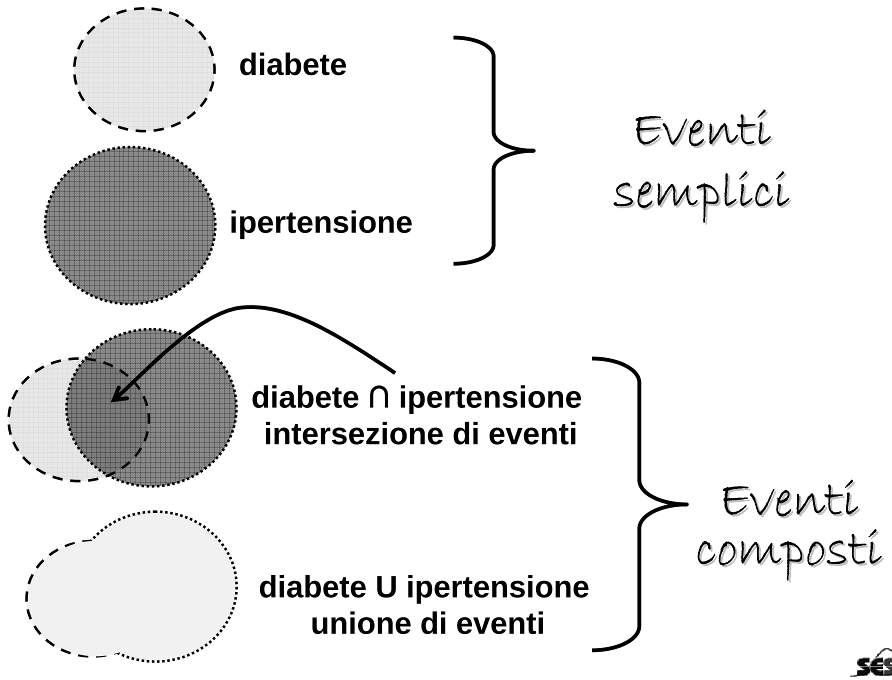
$C = A \cap B$

Esercizio: *nel lancio di un dado sia A = 'numero pari' e B = 'numero ≥ 4'; si determini l'insieme unione e l'insieme intersezione.*

$A = \{2, 4, 6\}$ **$B = \{4, 5, 6\}$**

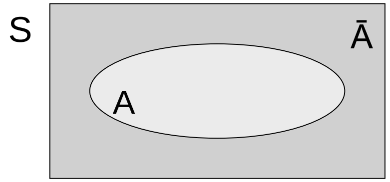
UNIONE → $A \cup B = \{2, 4, 5, 6\}$

INTERSEZIONE → $A \cap B = \{4, 6\}$



NEGAZIONE DI UN EVENTO: Ā

Dato un evento A, la sua negazione identifica un nuovo evento Ā costituito da tutti gli elementi di S non appartenenti ad A. Ā è detto **complemento di A in S**.



Segue che:
$A \cup \bar{A} = S$

Se due eventi A e B non hanno elementi in comune essi sono detti eventi **DISGIUNTI** o **MUTUAMENTE ESCLUSIVI** perché l'occorrenza dell'uno esclude l'altro.

Se A e B sono <u>mutuamente esclusivi</u> $\Rightarrow$ $A \cap B = \varnothing$

PROBABILITÀ

Lo spazio campionario S rappresenta l'insieme dei possibili risultati di un esperimento.

In genere, di fronte alla possibilità di un evento, esprimiamo la nostra **maggiore o minore fiducia** sul fatto che esso si verifichi

attribuiamo una maggiore o minore probabilità al verificarsi dell'evento

Esempi:

- *è <u>probabile</u> che oggi piova*
- *è <u>più probabile</u> che il tumore al polmone insorga in un fumatore che in un non fumatore*
- *è <u>più probabile</u> che il tumore al seno insorga in una donna che ha partorito per la prima volta dopo i 40 anni*

PROBABILITÀ

Il concetto di probabilità ci permette di graduare l'ambito delle possibilità o di precisare il grado di fiducia che abbiamo nel verificarsi di un evento.

La TEORIA DELLA PROBABILITA' ci permette di formulare delle **valutazioni numeriche di probabilità** e di ricondurle alle regole del calcolo matematico.

L'interpretazione di tali valori numerici dipende dal significato che viene attribuito al concetto di probabilità.

CONCEZIONE CLASSICA DELLA PROBABILITÀ

La probabilità di un evento A è il rapporto tra il numero di casi favorevoli al verificarsi di A (n) e il numero di casi possibili (N)

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

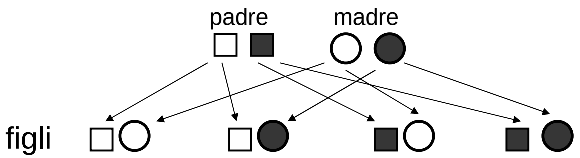
Esempi:

- **probabilità di estrarre un asso da un mazzo di 52 carte = 4/52 = 0.08**
- **probabilità di ottenere testa nel lancio di una moneta = 1/2 = 0.5**

- Tale definizione vale se i possibili risultati sono **equi-probabili** (gioco d'azzardo)
- Scarsamente applicabile a molte situazioni reali

Esempio di applicazione in medicina - Malattie genetiche

Se entrambi i genitori sono portatori sani del gene della talassemia o della fibrosi cistica, la probabilità di avere un figlio malato è una su quattro



CONCEZIONE FREQUENTISTA DELLA PROBABILITÀ

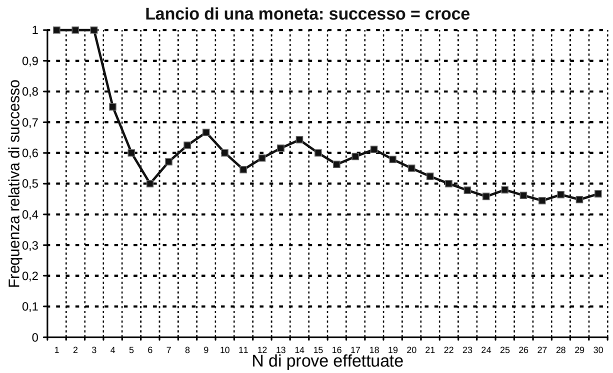
La probabilità di un evento A è la frequenza relativa di successo (occorrenza di A) in una serie **tendente all'infinito** di prove, ripetute **sotto identiche condizioni**:

$$P(A) = \lim_{N \to \infty} \frac{n}{N}$$

Nel caso della **concezione frequentista** la probabilità viene assegnata:

- **sulla base dei risultati di un esperimento ripetuto molte volte nelle stesse condizioni *(es. 1)***

Esempio 1: Frequenza dell'evento testa in una successione di lanci di una moneta.



Nel caso della **concezione frequentista** la probabilità viene assegnata:

- sulla base dei risultati di un esperimento ripetuto molte volte nelle stesse condizioni
- **sulla base di situazioni che possono essere ricondotte a tale contesto concettuale, ad esempio utilizzando statistiche correnti (*es. 2*).**

Esempio 2:

Prob. che un bambino italiano nasca morto nel 2000 =

$$= \frac{\textit{numero nati morti nel 2000}}{\textit{numero nati nel 2000}}$$

Non tutti gli eventi, pur valutabili in termini di probabilità, possiedono il requisito della **ripetitività sotto le stesse condizioni**

CONCEZIONE SOGGETTIVISTA DELLA PROBABILITÀ

La probabilità di un evento A è la **valutazione del grado di fiducia** che un individuo o un gruppo di individui può coerentemente formulare sull'occorrenza di A, **in base alle proprie opinioni e informazioni**

TEORIA BAYESIANA

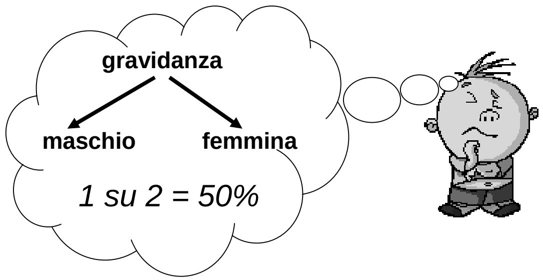
- Riguarda quei fenomeni per i quali l'**attesa** o la **convinzione** rispetto all'esito influisce sull'evento stesso (*interventi chirurgici; eventi che dipendono dalla propria volontà, capacità, ...*)
- Riguarda per lo più **eventi unici o irripetibili**

ESEMPIO: QUAL È LA PROBABILITÀ CHE UN NEONATO SIA FEMMINA?

L'ecografista, alla decima settimana di gravidanza, dice ai genitori che *80 su 100* il neonato è femmina (l'ecografista, secondo le sue opinioni ed informazioni, esprime coerentemente il suo grado di fiducia nell'avverarsi dell'evento "nascita di una femmina").

(definizione SOGGETTIVISTA di probabilità)

ESEMPIO: QUAL È LA PROBABILITÀ CHE UN NEONATO SIA FEMMINA?



(definizione CLASSICA di probabilità)

Però nel mondo, in assenza di interventi dell'uomo (aborti o infanticidi selettivi, omessa denuncia) nascono 1057 maschi ogni 1000 femmine.

1000 / (1000+1057) = 48,6%

(definizione FREQUENTISTA di probabilità)

TEORIA ASSIOMATICA DELLA PROBABILITÀ

Qualsiasi sia la definizione di probabilità, per probabilità (P) si intende una <u>funzione a valori reali</u> definita sullo spazio campionario S che soddisfa i seguenti **assiomi**:

1) per qualsiasi evento $A \subset S$, si ha che

$$0 \leq P(A) \leq 1$$

in particolare *P(A) = 1* se A è l'evento *certo*

P(A) = 0 se A è l'evento *impossibile*

2) $$P(S) = 1$$

TEORIA ASSIOMATICA DELLA PROBABILITÀ

3) se $\{A_1, A_2, \ldots, A_i, \ldots\}$ sono una sequenza finita o infinita di eventi mutuamente esclusivi (o disgiunti) di S, allora

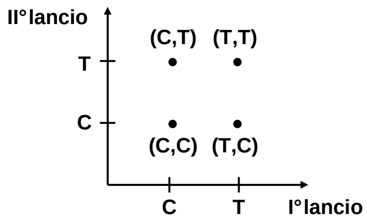
$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$$

Il modo più elementare per assegnare una funzione di probabilità allo spazio campionario S è quello di assegnare una probabilità ad ogni elemento di S

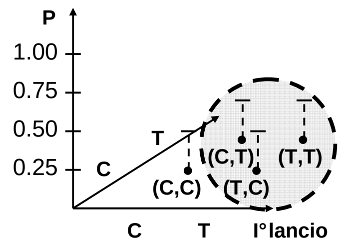
$\Rightarrow$ la probabilità corrispondente ad un qualsiasi **evento composto A** sarà definita come <u>somma delle probabilità degli eventi elementari contenuti in A</u> (*assioma 3*).

Esempio: Si consideri lo spazio campionario S associato al <u>lancio di una moneta per due volte consecutive</u>. In base alla definizione classica di probabilità, possiamo attribuire ad ogni punto dello spazio campionario S una probabilità P = 1/4 = 0.25.

SPAZIO CAMPIONARIO



FUNZIONE DI PROBABILITA'



Probabilità di ottenere almeno una testa = $P\{(T,C) \cup (C,T) \cup (T,T)\}$ = P(T,C) + P(C,T) + P(T,T) = 0.25 + 0.25 + 0.25 = **0.75**

MAZZO DI 52 CARTE

ESERCIZIO:

1. **calcolare la probabilità di estrarre <u>l'asso di cuori</u>**

   1 / 52 = 0.02

2. **calcolare la probabilità di estrarre <u>una carta rossa</u>**

   26 / 52 = 0.5

3. **calcolare la probabilità di estrarre <u>una figura</u>**

   12 / 52 = 0.23

4. **calcolare la probabilità di estrarre <u>un K</u>**

   4 / 52 = 0.08

SESM

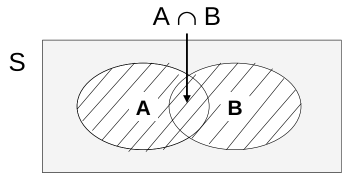
REGOLE DEL CALCOLO DELLA PROBABILITÀ

Il calcolo della probabilità è estremamente utile per stabilire sia la probabilità associata ad un evento, sia la probabilità associata ad un insieme di eventi.

REGOLA DELL'ADDIZIONE

1. Se A e B sono due eventi in S tali che $A \cap B \neq \varnothing$ (**eventi non disgiunti**):

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

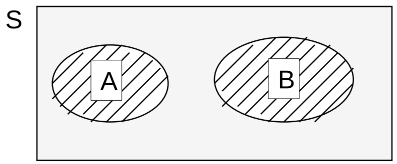


SESM

REGOLA DELL'ADDIZIONE

2. Se A e B sono due eventi in S tali che $A \cap B = \varnothing$ (**eventi disgiunti**):

$$P(A \cup B) = P(A) + P(B)$$



SESM

Esercizio: calcolare la probabilità di estrarre una carta rossa o una figura da un mazzo di 52 carte (eventi non disgiunti)

P(carta rossa) = 26 / 52 = ½ = 0.5

P(figura) = 12 / 52 = 0.23

P(carta rossa ∩ figura) = 6 / 52 = 0.11

P(carta rossa ∪ figura) =
= 0.5 + 0.23 -0.11 = 0.62

Esercizio: calcolare la probabilità di estrarre una figura o una carta compresa tra 3 e 6 da un mazzo di 52 carte (eventi disgiunti)

P(figura) = 12 / 52 = 0.23

P(carta 3-6) = 16 / 52 = 0.31

P(carta rossa ∩ figura) = 0 !!!

P(carta rossa ∪ carta 3-6) =
= 0.23 + 0.31 = 0.54

SESM

Se $\bar{A}$ è il complemento di A in S:

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

$$P(\bar{A}) = 1 - P(A)$$

Esercizio: se la probabilità di morire nel I° anno dal la diagnosi per un paziente affetto da tumore al polmone è pari a 0.30, qual è la probabilità di sopravvivere al I° anno?

P(sopravvivere) = 1 – 0.30 = 0.70

PROBABILITÀ CONDIZIONALE E REGOLA DELLA MOLTIPLICAZIONE

Talvolta è molto utile conoscere la probabilità di un evento A in S quando si è verificato un altro evento B in S $\rightarrow$ **PROBABILITA' CONDIZIONALE**

Esempio:

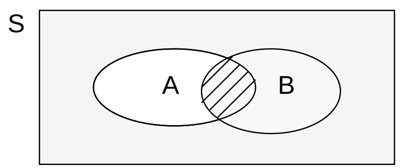
probabilità di uscita del 7 di quadri dato che è uscita una carta di quadri

probabilità di avere un tumore al polmone dato che si fuma

probabilità di avere il colera data la presenza di una gastroenterite acuta

Se A e B sono due eventi dello spazio campionario S, si definisce **probabilità condizionale di A dato B**:

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



N.B.: *lo spazio campionario dell'evento B diviene il nuovo spazio campionario*.

Dalla definizione di probabilità condizionale segue la **REGOLA DELLA MOLTIPLICAZIONE**:

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

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

Se il verificarsi di B non condiziona la probabilità del verificarsi di A, segue che:

$$P(A / B) = P(A)$$

e i due **eventi** sono detti **indipendenti**, ovvero:

$$P(A \cap B) = P(A) \bullet P(B)$$

INDIPENDENZA

Due eventi **A** e **B** si dicono **indipendenti** se e solo se

$$P(A \cap B) = P(A) \times P(B)$$

ovvero: *P(A|B) = P(A) e P(B|A) = P(B)*

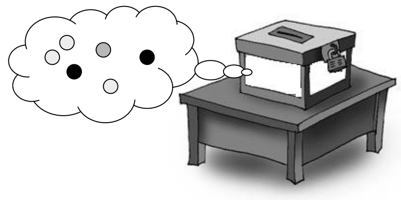
Se due eventi sono indipendenti il verificarsi dell'uno non influenza la probabilità di verificarsi dell'altro

esempio: elevati *livelli di glicemia* sono indipendenti dalla presenza di *ulcera*

SESM

Esempio:

Qual è la probabilità di estrarre senza reimbussolamento ***due*** *palline gialle da un'urna che contiene tre palline gialle, due nere e una verde?*



A = estraz. Iª pallina gialla → $P(A) = 3/6$

B = estraz. IIª pallina gialla → $P(B|A) = 2/5$

$$P(A \cap B) = P(A) \bullet P(B|A) = (3/6)(2/5) = \mathbf{1/5}$$

Se l'estrazione fosse con reimbussolamento:

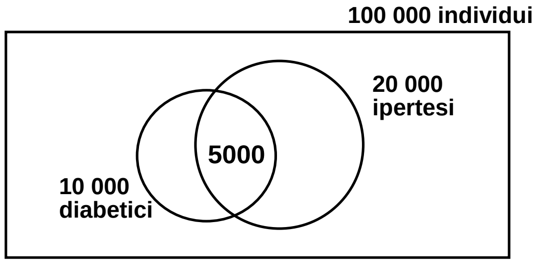
$$P(B|A) = P(B) = P(A) = 3/6$$

$$P(A \cap B) = (3/6)(3/6) = \mathbf{1/4}$$

SESM

ESERCIZIO: CALCOLO DELLE PROBABILITA'

In una popolazione di **100 000 individui** vi sono:
10 000 diabetici (e 90 000 non-diabetici)
20 000 ipertesi (e 80 000 non-ipertesi).
5000 persone che hanno sia il diabete che l'ipertensione

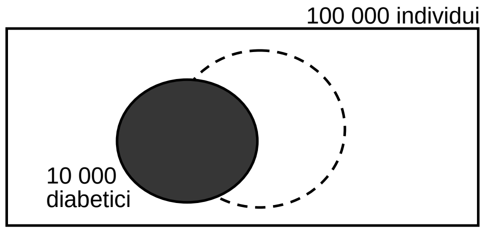


Qual è la probabilità di avere il diabete in quella popolazione?

Qual è la probabilità di avere l'ipertensione in quella popolazione?

SESM

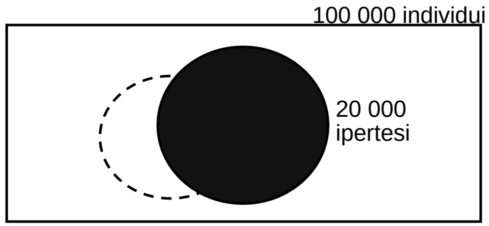
Qual è la probabilità di avere il diabete in quella popolazione?



N.B.
E' stato usato l'approccio frequentista: la probabilità è stata stimata dalla frequenza relativa

p (diabete) = 10 000 / 100 000 = 0,1 = 10%

Qual è la probabilità di avere l'ipertensione in quella popolazione?

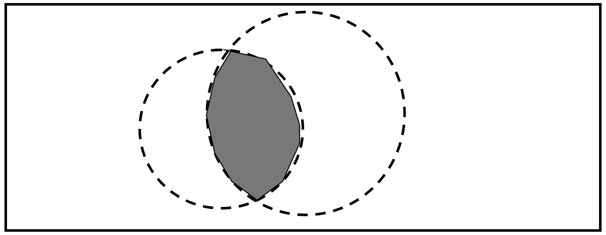


p (ipertensione) = 20 000 / 100 000 = 0,2 = 20%

SESM

Qual è la probabilità di avere il diabete **e** l'ipertensione (**sia il diabete che l'ipertensione**)?

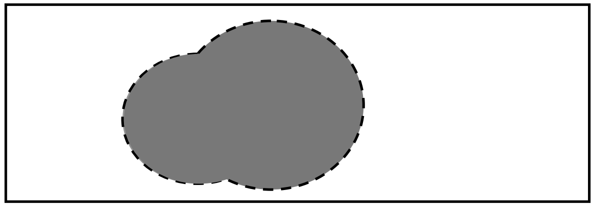
100 000 individui



p (diabete ∩ ipertensione) = 5 000 / 100000 = 0,05 = 5%

Qual è la probabilità di avere il diabete **o** l'ipertensione (**solo il diabete o solo l'ipertensione o entrambi**)?

100 000 individui



p(diabete U ipertensione) = (10000+20000-5000)/100000 = 25000/100000 = 0,25 = 25%

p(diabete U ipertensione) = p(diabete) + p(ipertensione) - p(diabete ∩ ipertensione) = 10% + 20% - 5% = 25%

Dipendenza e indipendenza statistica: rappresentazione grafica mediante diagramma di Venn

maschi **femmine**

ca. prostata **Ca. utero**

malaria **Talassemia**

Dipendenza e indipendenza statistica: rappresentazione grafica mediante diagramma di Venn

Psoriasi **Segno Gemelli**

HIV+ **Epatite B**

CALCOLO DEGLI ATTESI

Definita la probabilità di un evento o di una qualsiasi combinazione di eventi, è immediato definire il **numero di eventi attesi** in una serie di prove ripetute in modo casuale.

DEFINIZIONE: Se P(A) è la probabilità di comparsa dell'evento A, il **numero di eventi attesi E** in una serie di N prove sarà dato da:

$$\boxed{E = P(A) \bullet N}$$

NB: tanto più è grande N, tanto più il numero di successi osservati tenderà ad E

Esempio: *Se la probabilità di influenza per gli adulti tra i 20 e i 40 anni in un determinato mese dell'anno è **0.20**, quanti malati ci attendiamo in un campione casuale di **800** persone?*

E = 0.20 • 800 = 160

(numero atteso di soggetti con l'influenza)

Esempio: L*a probabilità di avere gli occhi azzurri è **0.15** e quella di possedere una FIAT è **0.25**. Sapendo che* P(avere occhi azzurri & avere FIAT) = 0.01*, calcolare quante persone in un campione di **1000** unità avranno gli occhi azzurri* <u>*oppure*</u> *una FIAT?*

"*avere gli occhi azzurri*" (A) e "*avere una FIAT*" (B) sono due eventi <u>non disgiunti</u>:

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

$= 0.25 + 0.15 - 0.01 = 0.39$

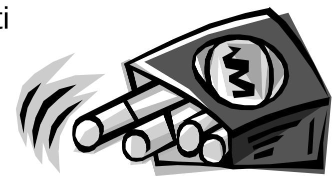
E = 0.39 • 1000 = 390

Esempio: *Si supponga che la probabilità di fumare sia **0.30** e quella di avere un tumore al polmone sia **0.01**. In un'indagine condotta su **1000** individui, sono stati osservati **15** soggetti fumatori con tumore al polmone. Secondo voi, i dati dimostrano che avere il tumore al polmone e fumare sono eventi indipendenti?*

E = (0.30 • 0.01) • 1000 = 3
(numero atteso di fumatori malati sotto l'ipotesi di indipendenza)

O = 15 (numero osservato di malati fumatori)

$\rightarrow$ **eccesso di 12 casi**



Prodotto di probabilità e sindrome plurimetabolica

Nello studio di Brunico *(Bonora et al, Diabetes 47: 1643-1649, 1998):*

N = 888

	Prevalenza
ridotta tolleranza glucidica	16,6%
dislipidemia	29,2%
iperuricemia	15,4%
ipertensione	37,3%

Se queste condizioni fossero indipendenti, la probabilità dell'intersezione (avere tutti e 4 i disturbi simultaneamente) sarebbe pari a:

0,166*0,292*0,154*0,373 = 0,0028 = **0,28%**

SESM

Prodotto di probabilità e sindrome plurimetabolica

p (avere tutte le patologie) = **0,28%**

Gli attesi (= soggetti con tutte e 4 le malattie sotto l'ipotesi di indipendenza) dovrebbero essere:

N * p = 888*0,0028 = **2,5**.

Invece se ne osservano 21!

Dal momento che gli osservati (21) sono molti di più degli attesi (2,5) si conclude che queste patologie non si riscontrano per caso negli stessi soggetti, ma rappresentano le diverse espressioni di una stessa patologia, la sindrome plurimetabolica.

SESM

STIME DI PROBABILITA'

EVENTI DI MAGGIORE INTERESSE IN AMBITO MEDICO:

- malattia / morte (M^+ , M^-)
- esposizione pregressa (E^+ , E^-)

PROBABILITA' DI MAGGIORE INTERESSE IN AMBITO MEDICO:

- $P (M^+ | E^+)$
- $P (M^+ | E^-)$

$$RR = \frac{P (M^+ | E^+)}{P (M^+ | E^-)}$$

SESM

esempio: la probabilità di morte per carcinoma polmonare tra gli individui di sesso maschile di età compresa tra i 55 e i 75 anni è:

$P (M+) = 0.02$ **informazione quantitativa di natura descrittiva**

esempio: la probabilità di morte per carcinoma polmonare tra gli individui di sesso maschile di età compresa tra i 55 e i 75 anni **fumatori / non fumatori** è:

$P (M+|F+) = 0.08$

$P (M+|F-) = 0.01$

informazione quantitativa sull'associazione tra esposizione e malattia

$$RR = \frac{0.08}{0.01} = 8$$

SESM

Le probabilità condizionali e le misure ad esse associate richiedono la stima delle probabilità associate ai punti dello SPAZIO CAMPIONARIO:

$$S = \{M^+E^+, M^+E^-, M^-E^+, M^-E^-\}$$

MALATTIA

ESPOSIZIONE	+	-
+	$M^+ \cap E^+$	$M^- \cap E^+$
-	$M^+ \cap E^-$	$M^+ \cap E^-$

MALATTIA

ESPOSIZIONE	+	-
+	$M^+ \cap E^+$	$M^- \cap E^+$
-	$M^+ \cap E^-$	$M^+ \cap E^-$

In assenza di informazioni a priori la probabilità dell'evento indicato in ciascuna delle 4 celle della tabella, P(…), può essere stimata **ripetendo l'esperimento** (<u>misurazione di esposizione e malattia su un individuo</u>) **n volte** e contando la frequenza n(…) con cui l'evento si verifica

$$P(\ldots) = \frac{n(\ldots)}{n}$$

Esempio: *valutiamo la relazione tra allattamento al seno (As) e insorgenza di infezioni del primo tratto respiratorio nei primi 4 mesi dalla nascita (IR)*

Indagine condotta sui nati tra il 1982 e il 1983 in una clinica ostetrica dell'Arizona

Infezione respiratoria

Allattamento al seno	+	-	
+	34	72	106
-	207	238	445
	241	310	551

FREQUENZE ASSOLUTE

As / **IR**	+	-	
+	34	72	106
-	207	238	445
	241	310	551

$$P(\ldots) = \frac{n(\ldots)}{n}$$

As / **IR**	+	-	
+	0.06	0.13	0.19
-	0.38	0.43	0.81
	0.44	0.56	1.00

STIME DI PROBABILITA'

Per il calcolo delle probabilità rilevanti si può indifferentemente utilizzare la tabella delle <u>frequenze assolute</u> o quella delle <u>stime di probabilità</u>

esempio: stimate la P (IR^+) nei primi 4 mesi di vita

		IR +	IR -	
As	+	34	72	106
	-	207	238	445
		241	310	**551**

		IR +	IR -	
As	+	**0.06**	0.13	0.19
	-	**0.38**	0.43	0.81
		0.44	0.56	1.00

$$P(IR^+) = \frac{\mathbf{241}}{\mathbf{551}} = 0.44$$

$P(IR^+) = P(IR^+ \cap E^+) + P(IR^+ \cap E^-) = \mathbf{0.06} + \mathbf{0.38} = \mathbf{0.44}$

esempio: stimate la P ($IR^+ \cup E^+$) nei primi 4 mesi di vita

		IR +	IR -	
As	+	**34**	72	**106**
	-	207	238	445
		241	310	**551**

		IR +	IR -	
As	+	**0.06**	0.13	**0.19**
	-	0.38	0.43	0.81
		0.44	0.56	1.00

$$P(IR^+ \cup E^+) = \frac{\mathbf{241 + 106 - 34}}{\mathbf{551}} = 0.57$$

$P(IR^+ \cup E^+) = P(IR^+) + P(E^+) – P(IR^+ \cap E^+) = \mathbf{0.44} + \mathbf{0.19} – \mathbf{0.06} =$
$= 0.57$

esercizio:

1. *stimate P(IR+ | E+)*
2. *stimate P(IR+ | E-)*
3. *calcolate il RR*
4. *stimate il numero di infezioni attese tra i bambini allattati al seno ($IR^+ \cap E^+$) assumendo l'indipendenza tra gli eventi*

		IR +	IR -	
As	+	**34**	72	**106**
	-	207	238	445
		241	310	**551**

		IR +	IR -	
As	+	**0.06**	0.13	**0.19**
	-	0.38	0.43	0.81
		0.44	0.56	1.00

SOLUZIONE

1. P(M+ **|** E+) = P(M+ $\cap$ E+) / P(E+) = (34/551) / (106/551) = 34 / 106 = 0.32
2. P(M+ **|** E–) = 207 / 445 = 0.47
3. RR = P(M+ **|** E+) / P(M+ **|** E–) = 0.32 / 0.47 = 0.68
4. num. atteso = P(M+ $\cap$ E+) • n = $\boxed{\text{P(M+) • P(E+)}}$ • n =
   = 0.44 • 0.19 • 551 = 46

ipotesi di indipendenza!

ESERCIZIO:

Nella tabella seguente è riportata la distribuzione di frequenza congiunta del sesso e della capacità vitale forzata (FVC) in cl:

		SESSO		
		Maschi	**Femmine**	**TOTALE**
FVC	**[200-300]**	**0**	**5**	**5**
	(300-400]	**4**	**27**	**31**
	(400-500]	**21**	**13**	**34**
	(500-600]	**20**	**1**	**21**
	(600-750]	**9**	**0**	**9**
	TOTALE	**54**	**46**	**100**

Qual è la probabilità che un soggetto abbia un valore dell'FVC > 500 cl?

Qual è la probabilità che un maschio abbia un valore dell'FVC > 500 cl?

Qual è la probabilità che un soggetto abbia un valore dell'FVC > 500 cl e sia femmina?

Qual è la probabilità che un soggetto sia femmina dato che ha un valore dell'FVC ≤ 400 cl?

Qual è la probabilità che un soggetto abbia un valore dell'FVC > 500 cl?

$$P(FVC > 500) = (21+9) / 100 = 0.30$$

Qual è la probabilità che un maschio abbia un valore dell'FVC > 500 cl?

$$P(FVC > 500 \mid maschio) = (20+9) / 54 = 0.54$$

Qual è la probabilità che un soggetto abbia un valore dell'FVC > 500 cl e sia femmina?

$$P(FVC > 500 \cap femmina) = (1+0) / 100 = 0.01$$

Qual è la probabilità che un soggetto sia femmina dato che ha un valore dell'FVC ≤ 400 cl?

$$P(femmina \mid FVC \leq 400) = (5+27) / (5+31) = 0.89$$