

Meta-Regressione Logistica, Modelli Makoviani, Reti Bayesiane e Metanalisi in MCMC and Gibbs Sampler



Verranno introdotti alcuni concetti fondamentali sull’Inferenza Bayesiana
e

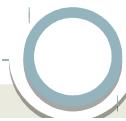
MCMC and Gibbs Sampler dopodiché, partendo da modelli molto semplici,

arriveremo a modelli di meta-regressione logistica sia con stime dirette dell’effect size sia con stime dirette ed indirette dell’effect size.

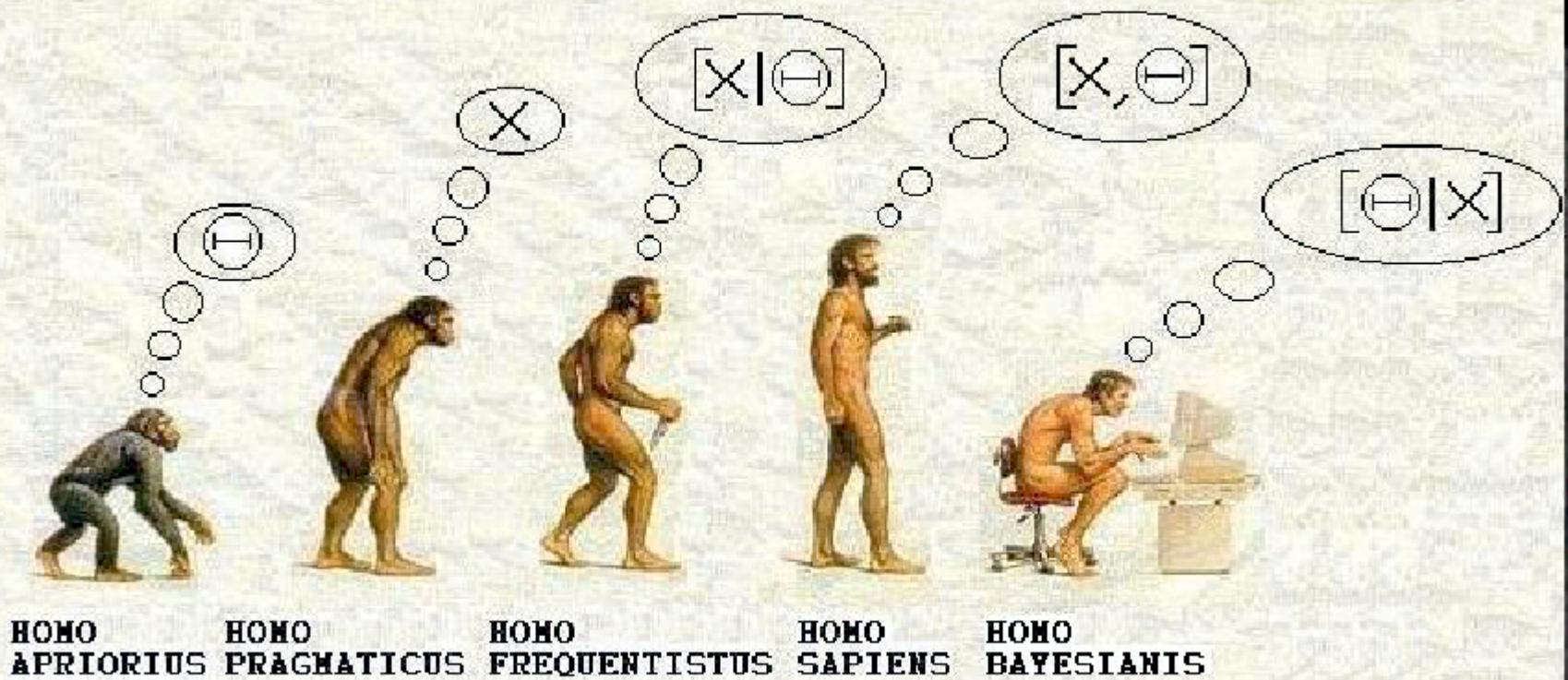
Come argomento successivo cercheremo di capire se è possibile applicare stime dirette e indirette di transizione su modelli markoviani in base all’effettività dei trattamenti. Il tutto applicando meta-analisi su studi eterogenei.

Il punto finale sarà capire se è possibile fare inferenza causale su reti Bayesiane e come utilizzare a tal scopo le meta-analisi dei vari studi - omici e clinici.

Introduzione



(YET ANOTHER) HISTORY OF LIFE AS WE KNOW IT...

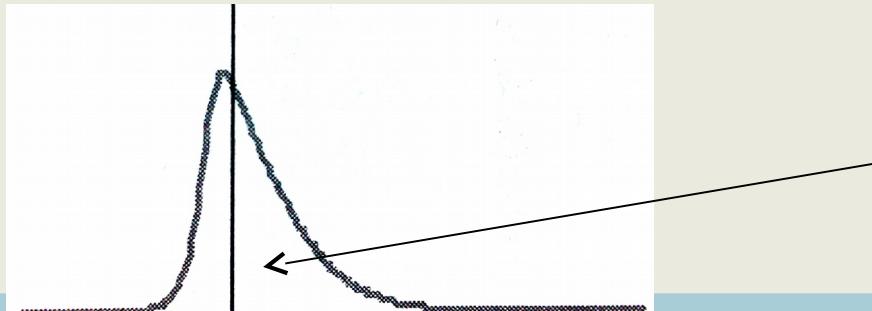


Il sole è sorto e tramontato per miliardi di anni. Il sole è tramontato anche stanotte. Con un'elevata probabilità, il sole domani sorgerà.

Il problema computazionale degli anni '80



- Con distribuzioni a priori coniugate non è quasi mai possibile ottenere in forma esplicita la distribuzione di probabilità finale di un parametro di interesse.
- Quello che si faceva era utilizzare approssimazioni analitiche e calcolo degli integrali con metodi basati sul teorema del limite centrale (approssimazione di Laplace, sviluppi in serie di Edgeworth e del punto di sella) per ottenere medie, mediane a posteriori.
- Figuriamoci strutturare un framework sulle meta-analisi e l'analisi dei costi dove risultava difficile stimare l'incremental net-benefit in problemi decisionali inerenti la salute pubblica:
 $E[INB(P_1, \delta, (U_1 - U_2), (C_1 - C_2))]$ nel continuo è un integrale.



La zona a destra dal punto di neutralità ci permette di attuare una buona scelta tra costi e salute

Markov Chain Monte Carlo



L'evoluzione è stata l'utilizzo di metodologie computazionali di Monte Carlo e Markov Chain.

Nel caso di una simulazione di Monte Carlo:

Supponiamo di voler ottenere delle informazioni dalla distribuzione a posteriori $p(\theta)$ bisognerebbe risolvere l'integrale I : con ad esempio $g(\theta) = \theta$ per la media :

$$I = \int_{\Theta} g(\theta)p(\theta)d\theta$$

Oppure con una simulazione di Monte Carlo posso ottenere M campioni i.i.d e così avere uno stimatore consistente :

Consistenza dalla Legge Forte dei Grandi Numeri:

$$\frac{X_1 + X_2 + \cdots + X_M}{M} \rightarrow \mu \text{ as } M \rightarrow \infty$$

Con \mathbf{X} v.a i.i.d e $E[X_i] = \mu$

$$\hat{I}_M = \frac{1}{M} \sum_{i=1}^M g(\theta^{(i)})$$

Supponiamo di voler effettuare un campionamento per la a posteriori $p(\theta|y)$ ma riusciamo a misurare campioni che sono solo **debolmente dipendenti**. Markov Chain descrive bene questa situazione:

$$p(\theta^{(t+1)}|\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}) = p(\theta^{(t+1)}|\theta^{(t)})$$

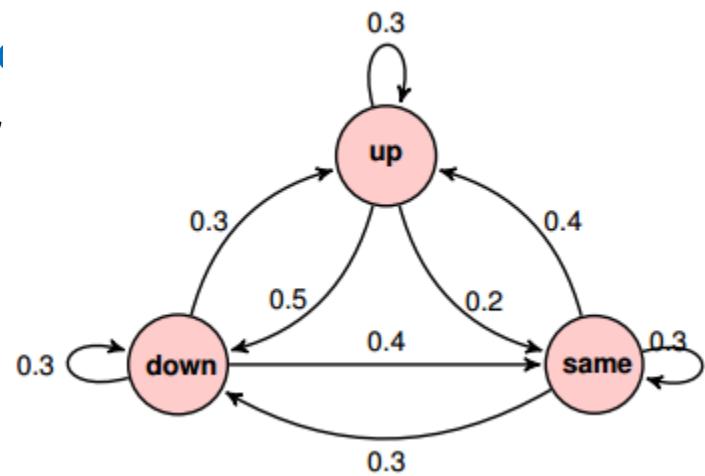
Per ogni parametro θ^t mi ritrovo nello stato θ^{t+1} con una certa probabilità $\pi^{t+1}(\theta) = \pi^{t-1}(\theta) \times \mathbf{P}$. In particolare sono interessato alla distribuzione stazionaria $\pi^*(\theta)$: (per ogni stato reversibilità: $\pi(x)P_t(x,y) = \pi(y)P_t(y,x)$) condizione sufficiente per la π^* .

Il mio obiettivo è costruire una catena di Markov tale che la sua $\pi^*(\theta)$ sia proprio $p(\theta|y)$. Avviando un algoritmo MCMC e portandolo alla convergenza ottengo una $\pi^*(\theta)$ con una realizzazione $(\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)})$. Da questa realizzazione ottengo dei campioni che ben approssimano $p(\theta|y)$ con Monte Carlo.

Dal **teorema ergodico per variabili debolmente** Markov gode di queste proprietà: aperiodica, ricorrente allora:

$$\frac{1}{M} \sum_{i=1}^M g(\theta_i) \rightarrow \int_{\Theta} g(\theta) \pi(\theta) d\theta$$

as $M \rightarrow \infty$, where π is the stationary distribution.



Metropolis-Hastings Algorithms



MH è un criterio generale per generare una catena di Markov che converge a $\pi(\mathbf{X})$ fissato

a priori. La generazione dello stato successivo X^t è un processo a due fasi: viene calcolato il rapporto di Metropolis ed in seguito si procede con l'accettazione o reiezione del valore in base a una procedura pseudo-

Algoritmo di Metropolis-Hastings (MH)

1. Al tempo 0, la catena si trova in $X_0 = x_0$
2. Al tempo t genera un valore *candidato* $Z = z \sim q(z|x_{t-1})$
3. Calcola il rapporto $\alpha(z, x_{t-1})$, dove

$$\alpha(z, x_{t-1}) = \min \left(\frac{q(x_{t-1}|z)\pi(z)}{q(z|x_{t-1})\pi(x_{t-1})}, 1 \right)$$

4. genera $\xi \sim U(0, 1)$ e poni

$$X_t = \begin{cases} z & \text{se } \xi \leq \alpha(z, x_{t-1}) \\ X_{t-1} & \text{se } \xi > \alpha(z, x_{t-1}) \end{cases}$$

MCMC and Gibbs Sampler



L'idea di fondo è di costruire la catena di Markov utilizzando, per generare i valori di ogni singola componente del vettore dei parametri, la distribuzione a posteriori di quella componente del parametro condizionata al valore assunto, nelle precedenti iterazioni, dalle altre componenti del parametro.

Parto da valori iniziali della catena e devo specificare qual'è la full condizionale sulla quale effettuare il campionamento: $\pi(\theta_i | \theta_{-i})$.

- al passo t -esimo genera

$$\theta_1^{(t)} \sim \pi(\theta_1 | \theta_2^{(t-1)}, \dots, \theta_k^{(t-1)})$$

$$\theta_2^{(t)} \sim \pi(\theta_2 | \theta_1^{(t)}, \dots, \theta_k^{(t-1)})$$

$$\theta_3^{(t)} \sim \pi(\theta_3 | \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_k^{(t-1)})$$

.....

$$\theta_k^{(t)} \sim \pi(\theta_k | \theta_1^{(t)}, \dots, \theta_{k-1}^{(t)})$$

Se non ho idea di che distribuzione abbia la congiunta utilizzo Metropolis-Hastings

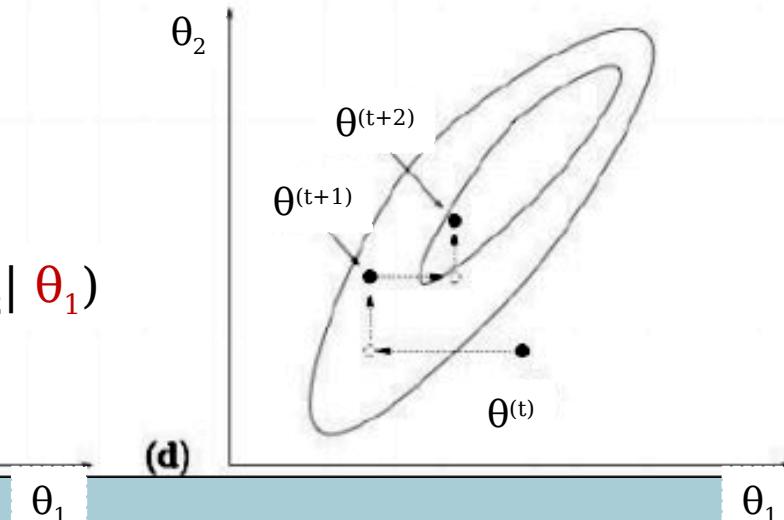
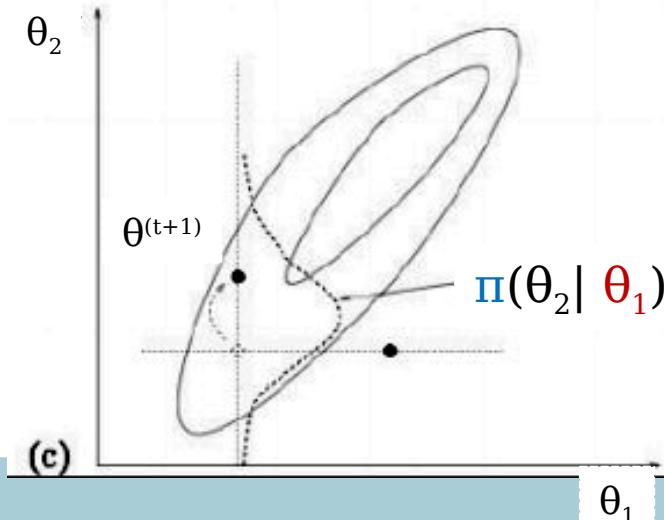
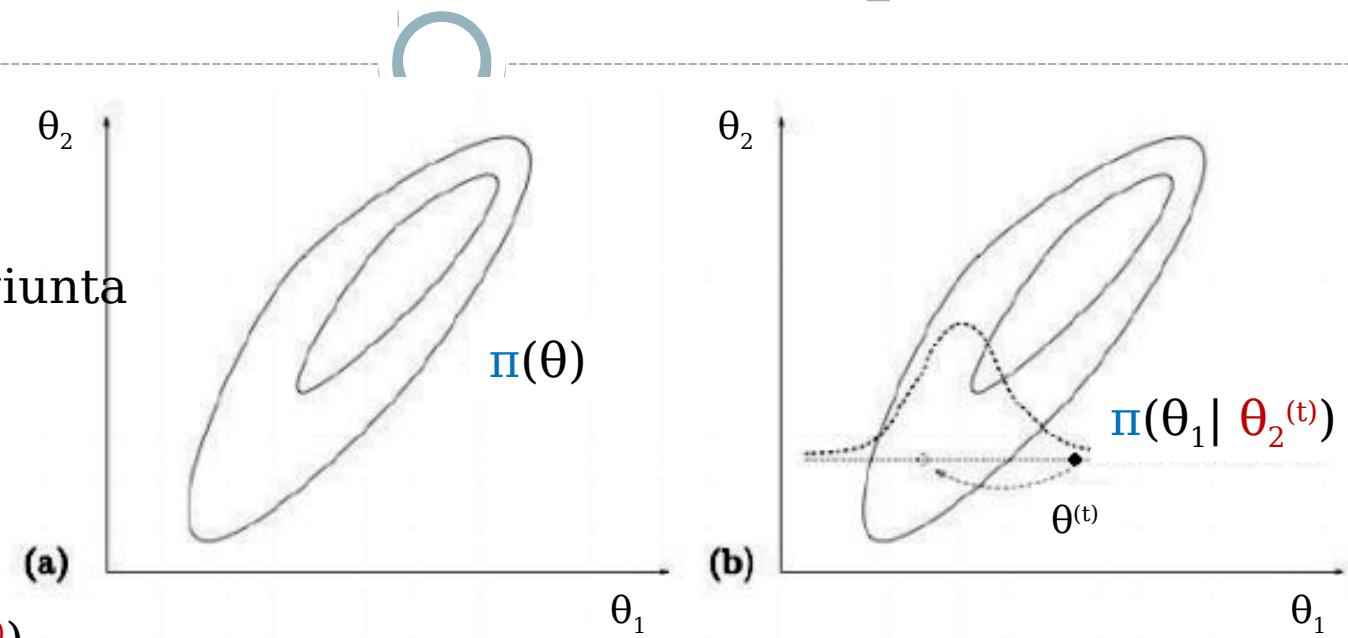
MCMC and Gibbs Sampler

$$(\theta_1, \theta_2) \sim \pi(\theta_1, \theta_2)$$

Se conosco la congiunta
posso ricavarmi la
full condizionale:

$$\theta_1^{(t+1)} \sim \pi(\theta_1 | \theta_2^{(t)}) \quad (\text{a})$$

$$\theta_2^{(t+1)} \sim \pi(\theta_2 | \theta_1^{(t+1)})$$



MCMC and Gibbs Sampler

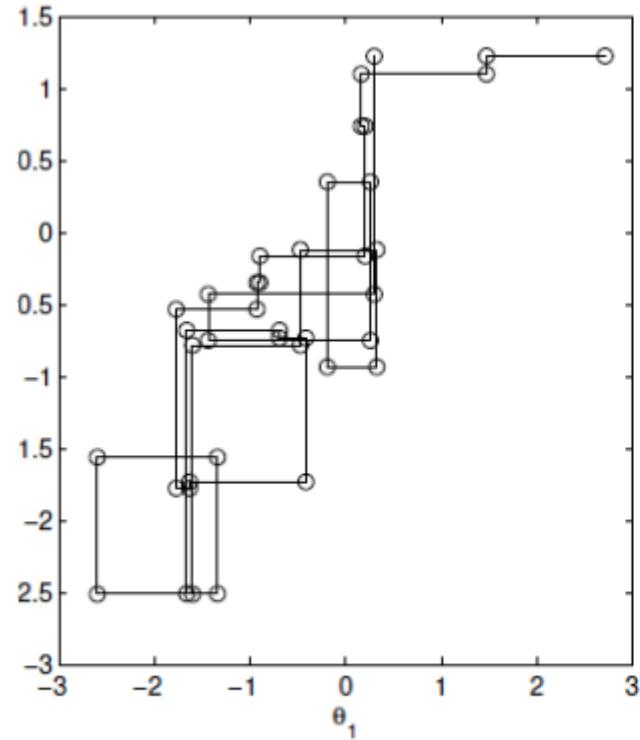


$(\theta_1, \theta_2) \sim \Pi(\theta_1, \theta_2)$

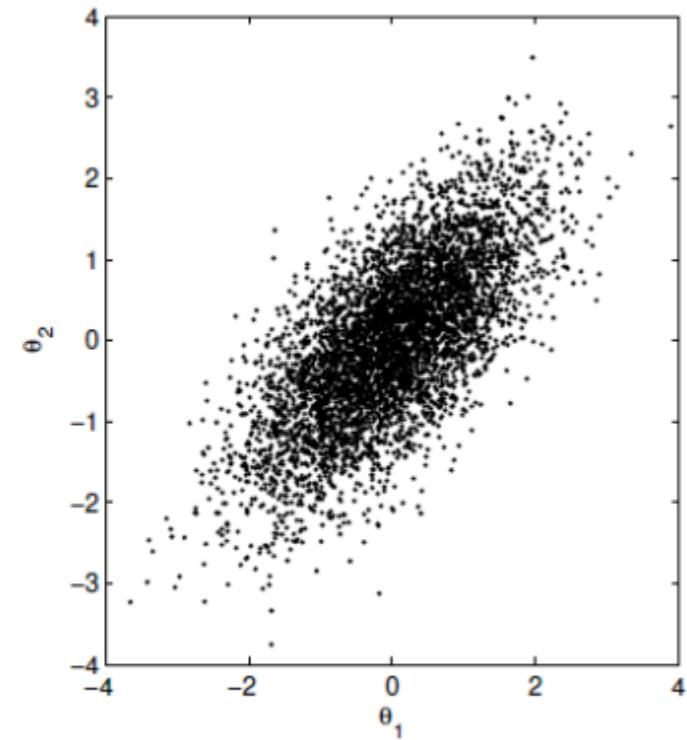
$\theta_1^{(t+1)} \sim \Pi(\theta_1^{(t)} | \theta_2^{(t)})$

$\theta_2^{(t+1)} \sim \Pi(\theta_2^{(t)} | \theta_1^{(t+1)})$

$\theta_1^{(t+1)}$



20 iterazioni



5000 iterazioni

Esempio di MCMC and Gibbs Sampler

Esempio 7.10 [Normale bivariata.]

Supponiamo di voler generare valori da una distribuzione normale bivariata, associata cioè ad una v.a. $\mathbf{Y} = (Y_1, Y_2)$ con media $(0, 0)$ e matrice di correlazione

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Da proprietà elementari della distribuzione normale sappiamo che le distribuzioni di ciascuna componente di \mathbf{Y} condizionata all'altra sono ancora normali; più precisamente,

$$Y_1 | Y_2 = u \sim N(\rho u, 1 - \rho^2)$$

$$Y_2 | Y_1 = v \sim N(\rho v, 1 - \rho^2)$$

È sufficiente allora scegliere un valore di partenza $\mathbf{y}^{(0)} = (y_1^{(0)}, y_2^{(0)})$ e generare una catena di Markov secondo il seguente schema

Gibbs sampling per Normale bivariata

- parti da $\mathbf{y}^{(0)} = (y_1^{(0)}, y_2^{(0)})$
- al passo t -esimo genera

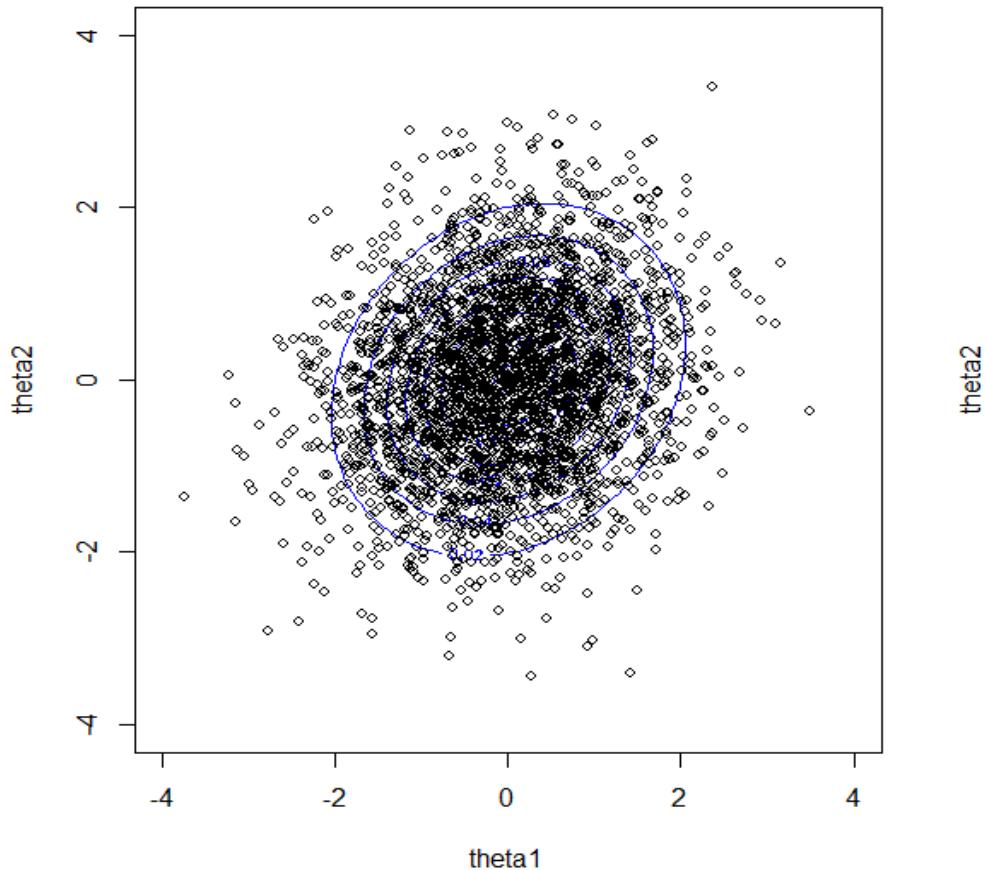
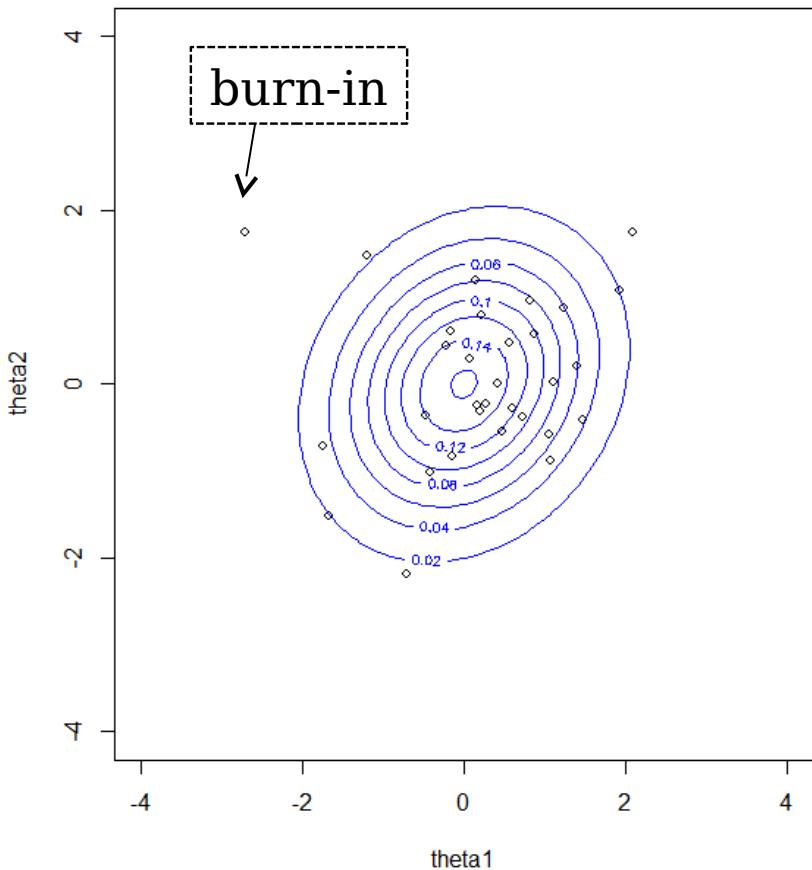
$$y_1^{(t)} \sim N(\rho y_2^{(t-1)}, 1 - \rho^2),$$

$$y_2^{(t)} \sim N(\rho y_1^{(t)}, 1 - \rho^2)$$

Conclusioni su MCMC e Gibbs Sampler



Dopo 100 e 1000 iterazioni di MCMC Gibbs Sampler in R:



La Misura Associativa: Odds Ratio



$$\frac{P(\text{Trattamento}/\text{NumCampioni})(1 - P(\text{Controllo}/\text{NumCampioni}))}{P(\text{Controllo}/\text{NumCampioni})(1 - P(\text{Trattamento}/\text{NumCampioni}))}$$

In epidemiologia l'**odds ratio** (OR) è uno degli indici utilizzati per definire il rapporto di **causa-effetto** tra due fattori.

In medicina l'**odds ratio** di un trattamento è il rapporto tra la frequenza con la quale un evento si verifica in un gruppo di pazienti e la frequenza con la quale lo stesso evento si verifica in un gruppo di pazienti *di controllo*.

OR < 1 il trattamento è efficace.

OR = 1 il trattamento è irrilevante.

OR > 1 il trattamento è inefficace.

Ottenere l'Effect Size Medio



Dai modelli che studieremo abbiamo come **obiettivo** è quello di ottenere un *effect size d* medio.

Nella Meta-Analisi è un indice che misura la grandezza di un evento, dimensione di un effetto.

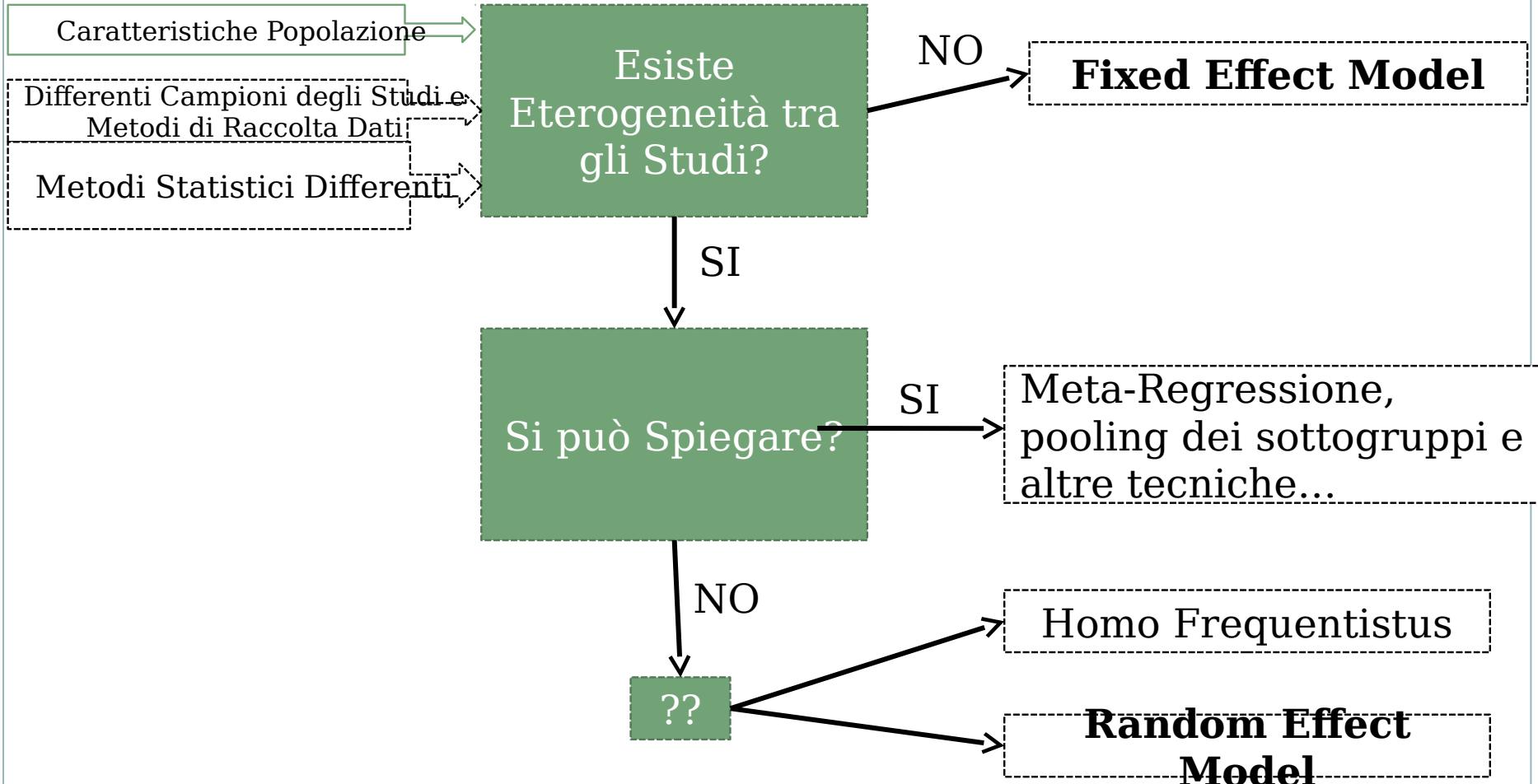
L'*effect size* può basarsi:

- 1) sulle medie (*raw unstandardized mean difference; standardized mean difference, responses ratios*);
- 2) sui dati binari (*risk ratio; odds ratio; risk difference*);
- 3) sulle correlazioni (r).

Randomize
d
Controlled
Trials

$$P(d \geq 0) = P(OR \geq 1)$$

L'Eterogeneità tra gli Studi c'è o non c'è?



Aspirine VS Placebo



Comparazione tra l'utilizzo di aspirine rispetto che placebo per prevenire la morte da infarto del miocardio.

7 Studi	Utilizzo di Aspirine (A)		Utilizzo di Placebo (B)	
	<i>Tot</i>	<i>Succes</i> <i>s</i>	<i>Tot</i>	<i>Succes</i> <i>s</i>
MRC-1	615	566	624	557
CDP	758	714	771	707
MRC-2	832	730	850	724
GASP	317	285	309	271
PARIS	810	725	406	354
AMIS	2267	2021	2257	2038
ISIS-2	8587	7017	8600	6880

$$OR = \frac{ad}{bc}$$

		Ln(OR)	Var(Ln(OR))
MRC-1	-0.3289	0.0389	
CDP	-0.3845	0.0412	
MRC-2	-0.2196	0.0205	
GASP	-0.2222	0.0648	
PARIS	-0.2255	0.0352	
AMIS	0.1246	0.0096	
ISIS-2	-0.1110	0.0015	



Fixed-effect model: modello statistico utilizzato per combinare gli effect sizes e, quindi, per calcolare l'effect size medio d .

Tale modello assume che gli studi inclusi nella meta-analisi abbiano lo stesso *effect size* (che pertanto è fisso).

Fixed Effect Model

```

model
{
  for (i in 1:Nstud)
  {
    P[i] <- 1/V[i]
    y[i] ~ dnorm(d, P[i])
  }
  d ~ dnorm(0, 1.0E-5)
  OR <- exp(d)
}

```

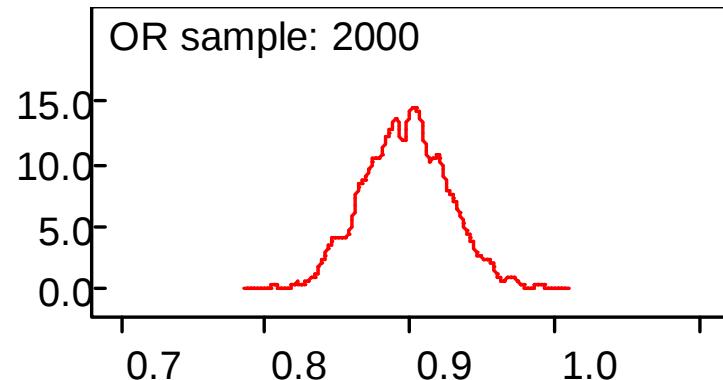
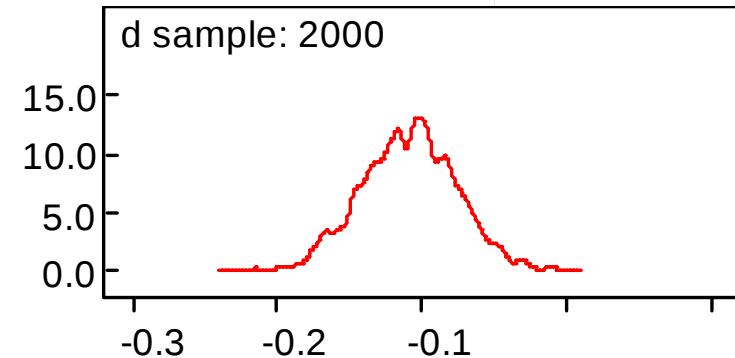
```

list(
y=c(-0.328, -0.384, -0.219, -0.22, -0.225,
0.124, -0.11),
V=c(.0388,.041,.020,.064,.035,.0096,.0015),
Nstud=7
)

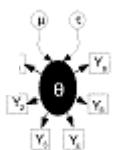
```

Catena Valore Iniziale: `list(d=0)`

$$Y_i \sim \text{Normal} \left(d, v^2_i \right) \quad i = 1, \dots, k$$



node	mean	sd	MC error	2.5%	median	95.0%	start	sample	
OR	0.8973	0.02981	6.545E-4	0.8409	0.8978	0.9466	0.9562	2001	
d	-0.1089	0.03323	7.278E-4	-0.1733	-0.1079	-0.05492	-0.04483	2001	



L'effect-size d è < 0 sta ad indicare che l'ampiezza dell'effetto è più rilevante per il trattamento rispetto al campione di controllo con un OR efficace. L'aspirina vince.

$$P(d \geq 0) = P(OR \geq 1)$$

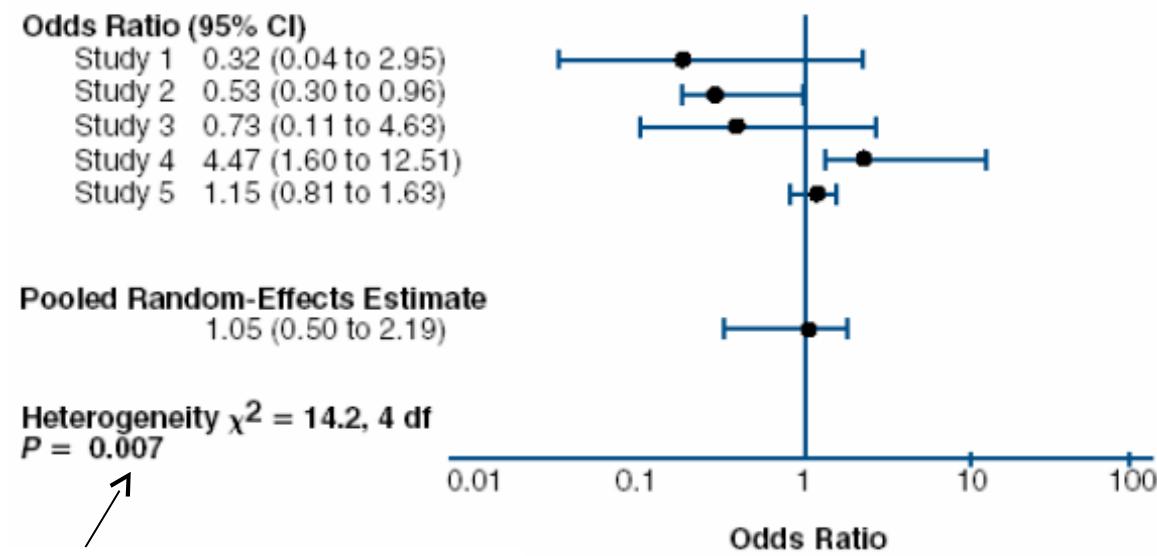
Random-effects model: modello statistico utilizzato per combinare gli *effect sizes* e, quindi, per calcolare l'*effect size* medio d .

Tale modello assume che gli *effect sizes* δ_i considerati nella meta-analisi possano variare tra di loro:

- 1) a causa della varianza *within-group*
- 2) a causa della reale variabilità (*between-study variance*) che esiste tra i parametri τ^2_i

Come trovare l'eterogeneità tra gli Studi?

Il test misura se le differenze tra i risultati dei singoli studi sono più grandi di quelli che ci si aspetterebbe se tutti gli studi avessero misurato lo stesso effetto e se le differenze osservate fossero dovute solamente al caso.



E' più che per caso...

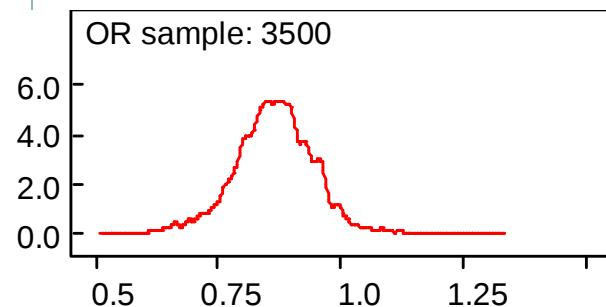
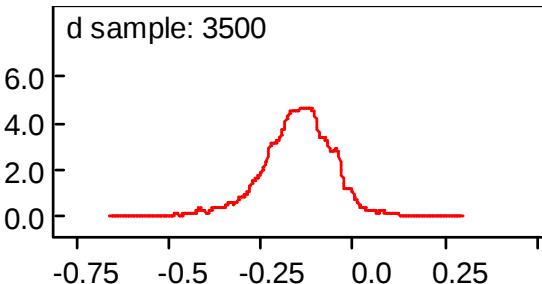
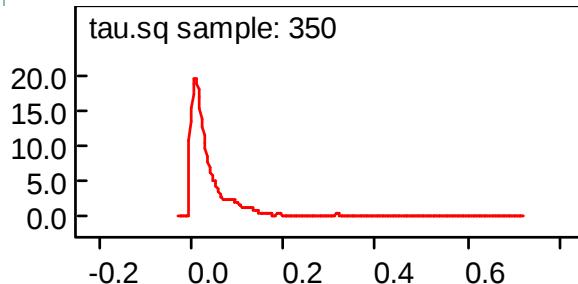
Random Effect Model

$$Y_i \sim \text{Normal}(\delta_i, v_i^2) \quad i = 1, \dots, k$$

$$\delta_i \sim \text{Normal}(d, \tau^2) \quad i = 1, \dots, k$$



Stima dell'effectsize Medio e Varianza



model

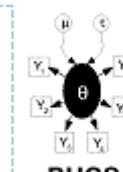
```
{
  for (i in 1:Nstud)
  {
    P[i] <- 1/V[i]
    y[i] ~ dnorm(delta[i], P[i])
  }
  delta[i] ~ dnorm(d, prec)
}
d ~ dnorm(0, 1.0E-5)
OR <- exp(d)
tau~dunif(0,10)
tau.sq<-tau*tau
prec<-1/(tau.sq)
}
```

Data: list(

```
y=c(-0.328, -0.384, -0.219, -0.22, -0.225,
0.124, -0.11),
V=c(.0388,.041,.020,.064,.035,.0096,.0015),
Nstud=7)
```

Initial Values: list(d=0, tau =1,
delta=c(0,0,0,0,0,0))

node	mean	sd	MC error	97.5%	start	sample
OR	0.8646	0.08405	0.003318	1.026501	3500	
d	-0.1503	0.0988	0.003922	0.02605	501	3500
prec	15380.0	344200.0	14260.0	8468.0	501	3500
tau.sq	0.0427	0.06585	0.002756	0.2039	501	3500



Regressione Logistica: Dicotomia

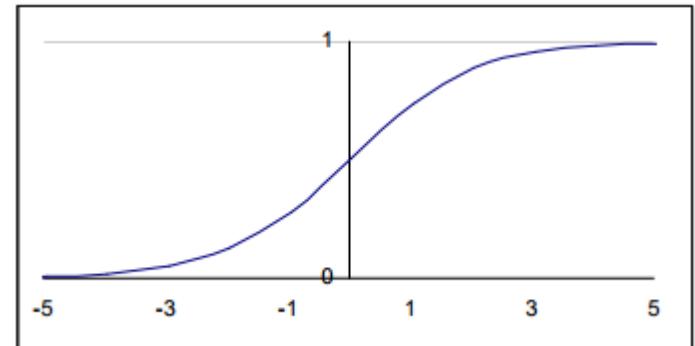
Si tratta di costruire un modello di regressione dicotomico per Y che può valere rispettivamente **0** all'assenza dell'attributo e **1** alla presenza dell'attributo.

Può essere utile nel distinguere un individuo sano\malato, o per un farmaco se ha effetto\non ha effetto.

In un modello di regressione la quantità che si ipotizza funzione di X è il valore medio aritmetico della variabile dipendente Y condizionato ad un dato x cioè $E[Y|x] = P(Y=1|x) = \pi(x)$, e quindi ancora un integrale nel continuo.

Il modello di regressione per Y è dunque: $Y = \pi(x) + \varepsilon$

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} = \frac{e^{\beta_0 + \sum_{j=1}^p \beta_j x_j}}{1 + e^{\beta_0 + \sum_{j=1}^p \beta_j x_j}}$$



In figura il caso più semplice con una sola variabile esplicativa X e $\beta_0 = 0$ e $\beta_1 = 1$

Regressione Logistica: OR e linearità



Si consideri, ora, la seguente funzione di $\pi(x)$, detta **logit**($\pi(x)$) :

$$\text{logit}(\pi(x)) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right];$$

Che è il **logaritmo naturale del rapporto della probabilità condizionata di possedere l'attributo e la probabilità condizionata di non possederlo**.

Il rapporto fra probabilità associate ad una dicotomia, cioè fra probabilità complementari, è detto odds ratio - OR.

Sostituendo è possibile dimostrare che:

$$\text{logit}(\pi(x)) = \beta_0 + \sum_{j=1}^p \beta_j x_j$$

Quindi mentre $\pi(x)$ è non lineare nelle variabili X il logaritmo dell'odds , detto **logit** , è funzione lineare delle stesse.

Regressione Logistica: Odds Ratio



$$P(Y = 1) = \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}}$$

funzione logistica

$$P(Y = 0) = 1 - \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}} = \frac{1}{1 + e^{\alpha + \beta X}}$$

$$odds_{Y=1} = \frac{e^{(\alpha + \beta X)}}{1 + e^{(\alpha + \beta X)}} = \frac{1}{\frac{1 + e^{(\alpha + \beta X)}}{e^{(\alpha + \beta X)}}} = e^{(\alpha + \beta X)}$$

$\ln(e^x) = x$

$$\ln(odds_{Y=1}) = \alpha + \beta X$$

Regressione Logistica: Due Strade



Nella maggior parte dei modelli di meta-regressione, l'unità dell'analisi, che è l'osservazione nel modello di regressione, è uno studio.

Il calcolo dei log odds ratio = Y = la variabile dipendente è stimata da una **intercetta** δ_i e dai **predittori associati** βx_i oppure soltanto dall'intercetta δ_i .

Vedremo entrambi i casi sul modello di Thompson et all:

$$Y_i \sim \text{Normal}(\delta_i, V_i) \quad Y_i \sim \text{Normal}(\delta_i + \beta x_i, V_i)$$

I predittori generalmente sono covariate collegate agli studi esempio **latitudine[i]**

La meta-regressione logistica punta alla stima degli effetti di un trattamento analizzando le differenze tra gli studi espressi in LOR, e determinando se c'è eterogeneità tra gli studi.

Esistono diversi tipi di meta-regressione noi utilizzeremo la random effects meta-regression.

Random Effect Model con Meta-Regressione Logistica

7 Studi	Utilizzo di Aspirine (A)		Utilizzo di Placebo (B)	
	Tot nA	Morti rA	Tot nB	Morti rB
MRC-1	615	49	624	67
CDP	758	44	771	64
MRC-2	832	102	850	126
GASP	317	32	309	38
PARIS	810	85	406	52
AMIS	2267	246	2257	219
ISIS-2	8587	1570	8600	1720

$rA_i \sim \text{Binomial}(pA_i, nA_i)$
 $rB_i \sim \text{Binomial}(pB_i, nB_i)$

$\text{logit}(pA_i) = \mu_i$
 $\text{logit}(pB_i) = \mu_i + \delta_i$
 $\delta_i \sim \text{Normal}(d, \tau)$

```

Model {
  for( i in 1 : k ) {
    rA[i] ~ dbin(pA[i], nA[i])
    rB[i] ~ dbin(pB[i], nB[i])
    logit(pA[i]) <- mu[i]
    logit(pB[i]) <- mu[i] + delta[i]
    mu[i] ~ dnorm(0.0,1.0E-5)
    delta[i] ~ dnorm(d, prec)
  }
  OR <- exp(d)
  d ~ dnorm(0.0,1.0E-6)
  tau~dunif(0,10)
  tau.sq<-tau*tau
  prec<-1/(tau.sq)
  delta[9] <- d
  delta[10] ~ dnorm(d, prec)
}

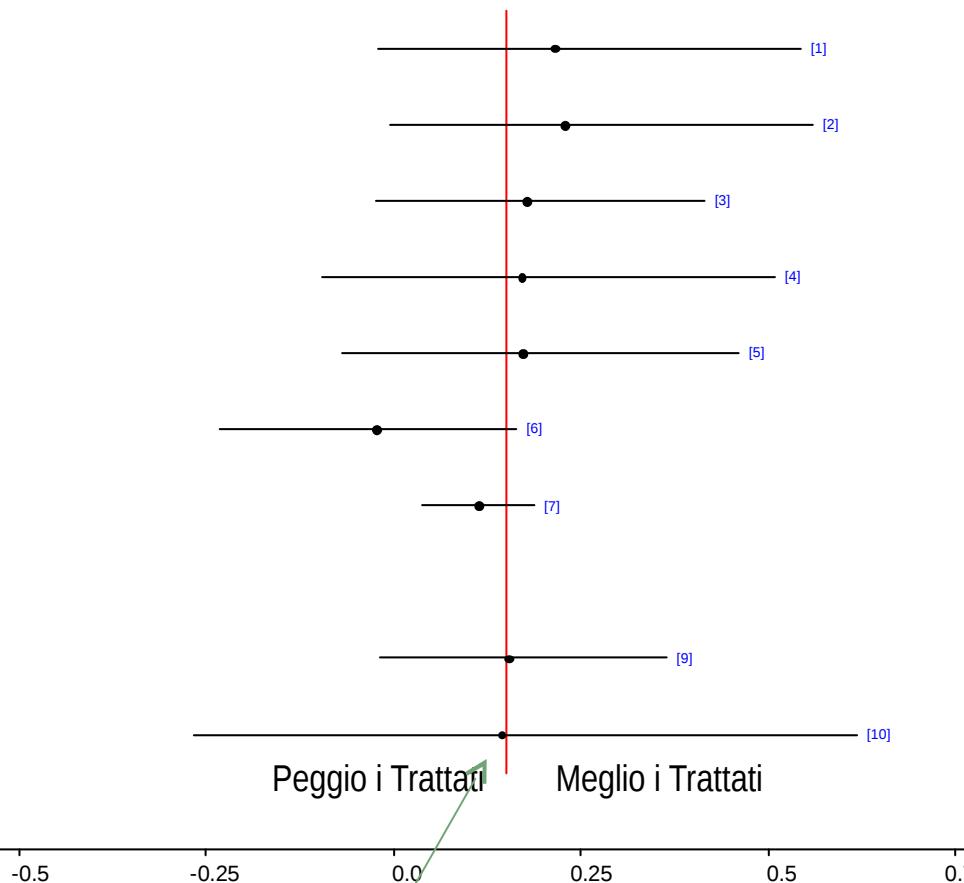
```

Data
list(
 rA=c(49,44,102,32,85,246,1570),
 rB=c(67,64,126,38,52,219,1720),
 nA=c(615,758,832,317,810,2267,8587),
 nB=c(624,771,850,309,406,2257,8600),
 k=7)

Initial Values
list(d=0, tau=2, delta=c(0,0,0,0,0,0,NA,NA,1),
mu=c(0,0,0,0,0,0,0))

Random Effect Model con Meta-Regressione Logistica

caterpillar plot: delta



Perché ci sono pochi dati.

```
Model {  
  for( i in 1 : k ) {  
    rA[i] ~ dbin(pA[i], nA[i])  
    rB[i] ~ dbin(pB[i], nB[i])  
    logit(pA[i]) <- mu[i]  
    logit(pB[i]) <- mu[i] + delta[i]  
    mu[i] ~ dnorm(0.0,1.0E-5)  
    delta[i] ~ dnorm(d, prec)  
  }  
  OR <- exp(d)  
  d ~ dnorm(0.0,1.0E-6)  
  tau~dunif(0,10)  
  tau.sq<-tau*tau  
  prec<-1/(tau.sq)  
  delta[9] <- d  
  delta[10] ~ dnorm(d, prec)  
}  
Data
```

```
list(  
  rA=c(49,44,102,32,85,246,1570),  
  rB=c(67,64,126,38,52,219,1720),  
  nA=c(615,758,832,317,810,2267,8587),  
  nB=c(624,771,850,309,406,2257,8600),  
  k=7)
```

Initial Values

```
list(d=0, tau=2, delta=c(0,0,0,0,0,0,NA,NA,1),  
mu=c(0,0,0,0,0,0,0))
```

Vaccino VS Nessun Vaccino - Tubercolosi

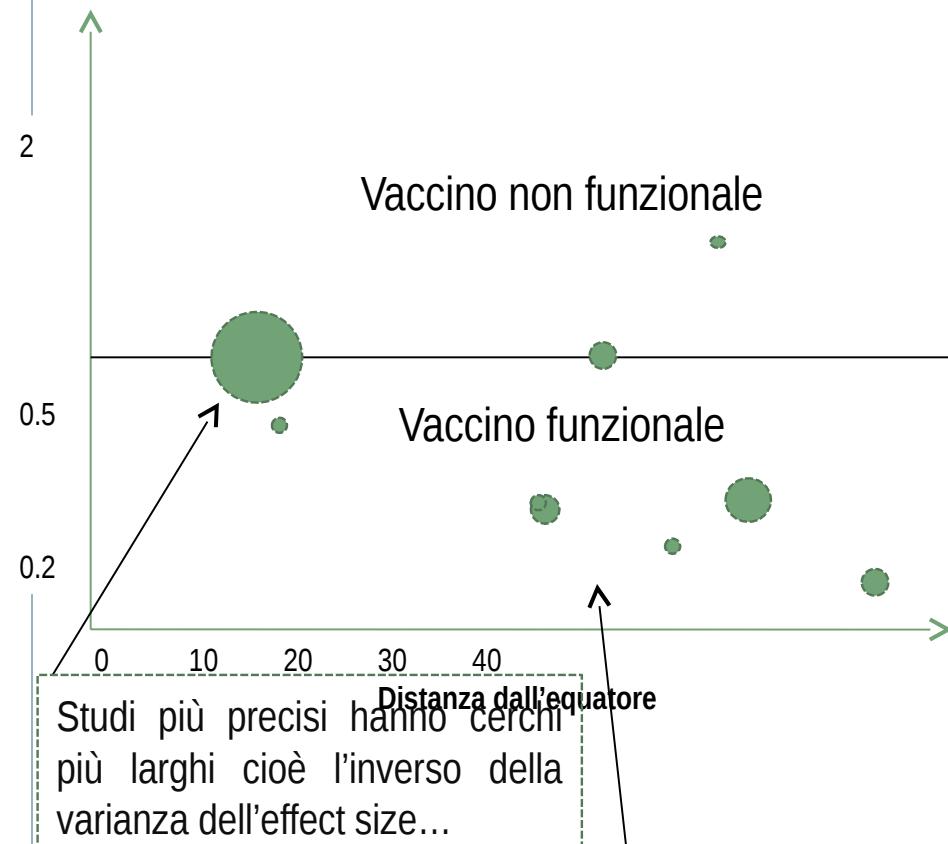
Studi	Latitudine	Vaccinati A		Non Vaccinati B	
		Malati rB	rB + Sani = nB	Malati rA	rA + Sani = nA
1	44	4	119	11	128
2	55	6	300	29	274
3	42	3	228	11	209
4	52	62	13.536	248	12.619
5	13	33	5.036	47	5.761
6	44	180	1.361	372	1.079
7	19	8	2.537	10	619
8	13	505	87.886	499	87.892
9	-27	29	7.470	45	7.232
10	42	17	1.699	65	1.600
11	18	186	50.448	141	27.197
12	33	5	2.493	3	23.338
13	33	27	18.886	29	17.825

Random Effect Model con Meta-Regressione Logistica

Utilizzo delle covariate

Inseriamo semplicemente il nostro coefficiente di regressione β moltiplicato per l' i-esimo predittore cioè la covariata del k-esimo studio

x_i :



$$Y_i \sim \text{Normal}(\delta_i + \beta x_i, V_i)$$

$$\delta_i \sim \text{Normal}(d, \tau^2)$$

$$\beta \sim \text{Normal}(0, 10^6)$$

model

```
{  
  for( i in 1 : Nstud ) {  
    rA[i] ~ dbin(pA[i], nA[i])  
    rB[i] ~ dbin(pB[i], nB[i])  
    logit(pA[i]) <- mu[i]  
    logit(pB[i]) <- mu[i] + delta[i] + beta*latitudine[i]  
    mu[i] ~ dnorm(0.0, 1.0E-6)  
    delta[i] ~ dnorm(d, prec)  
  }  
}
```

$$d \sim \text{dnorm}(0.0, 1.0E-6)$$

$$\tau \sim \text{unif}(0, 10)$$

$$\tau_{\text{sq}} \leftarrow \tau * \tau$$

$$\text{prec} \leftarrow -1 / (\tau_{\text{sq}})$$

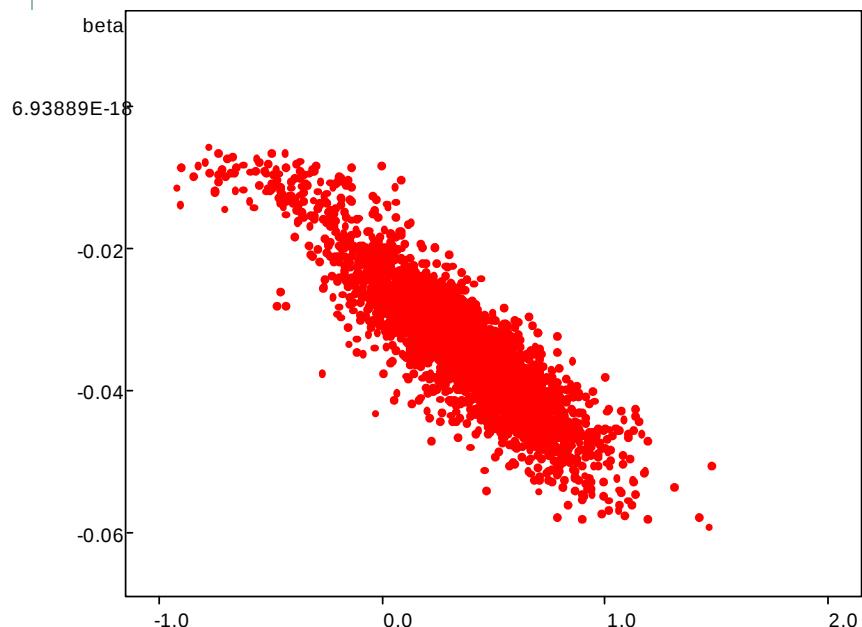
$$\beta \sim \text{dnorm}(0.0, 1.0E-6)$$

$$\text{OR} \leftarrow \exp(d)$$

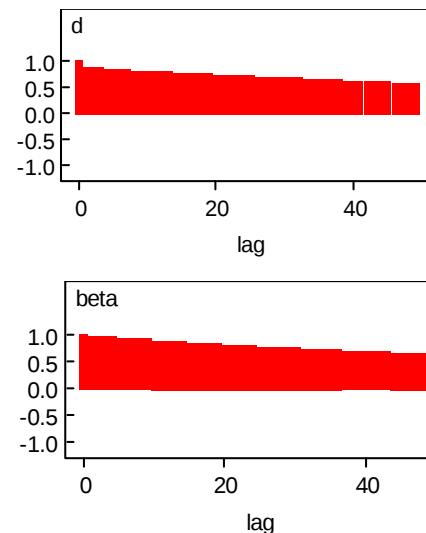
}

Beta = -0.03 (95% CrI -0.05 to -0.01) questo significa che i LOR decrescono linearmente, cioè all'allontanarsi dall'equatore il trattamento (vaccino per la tubercolosi) aumenta di size effect

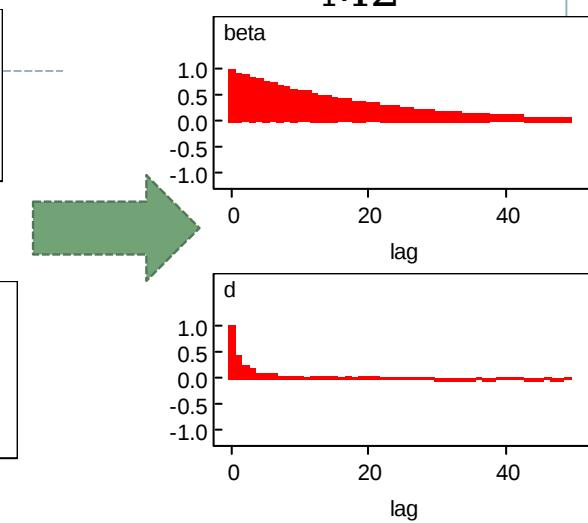
Random Effect Model con Meta-Regressione Logistica



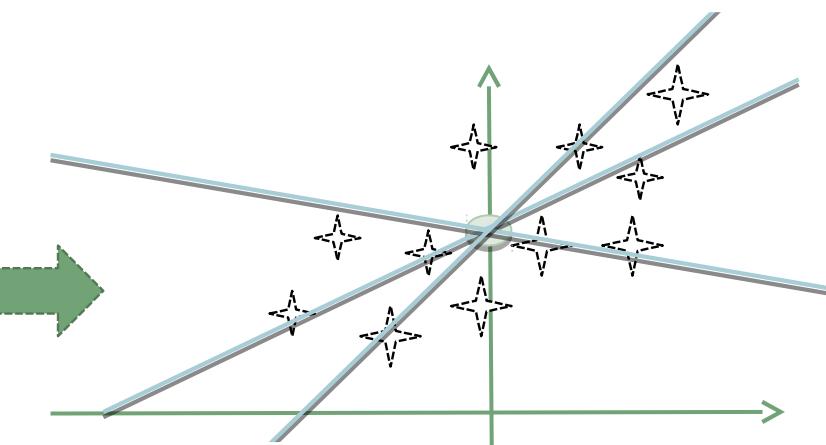
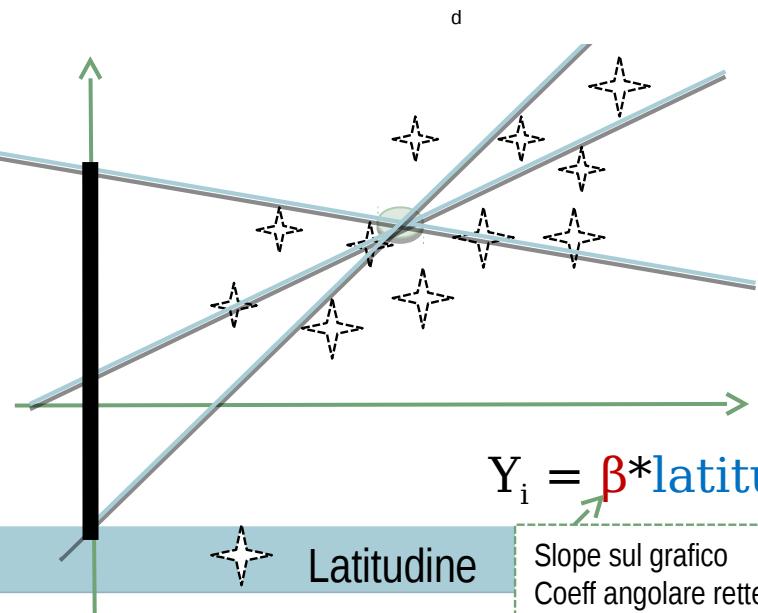
M1



M2



$$Y_i = \beta^*(\mu_i - \mu(\text{latitudine}_i)) + \mu_i + \delta_i$$



$$Y_i = \beta^* \text{latitudine}_i + \mu_i + \delta_i$$



Latitudine

Slope sul grafico
Coeff angolare rette blu

Intercetta area in nero.

Random Effect Model con Meta-Regressione

Logistica

Individual Patient Data



La meta-regressione con covariate del tipo **Individual Patient Data** (età, % white, dati omici e clinici individuali) è definita una **MetaAnalisi Debole**.[1]

I suoi risultati però possono essere utilizzati per risalire agli IPD suddividendo ad esempio i gruppi più vicini e più lontani dall'equatore e così aggiungendo caratteristiche agli stessi.

Modelli specifici per IPD con risultati nel continuo e ordinali esistono [2][3].

[1] Individual Patient-level and Study-level Meta-analysis for Investigating Modifiers of Treatment Effect - Satoshi Teramukai

[2] Meta-analysis of continuous outcome data from individual patients. Higgins

[3] Meta-analysis of ordinal outcome using patient data. Whitehead

Random Effect Model a MetaRegessione Logistica

BaseLine Risk

```
model
{
for( i in 1 : Nstud ) {
rA[i] ~ dbin(pA[i], nA[i])
rB[i] ~ dbin(pB[i], nB[i])
logit(pA[i]) <- mu[i]
//mean(latitudine[i]) = 3.7
logit(pB[i]) <- mu[i] + delta[i] + beta*(mu[i]-(-3.7))
mu[i] ~ dnorm(0.0,1.0E-3)
delta[i] ~ dnorm(d, prec)
temp[i] <- col_dif[i]
}
OR <- exp(d)
d ~ dnorm(0.0,1.0E-6)
tau~dunif(0,10)
tau.sq<-tau*tau
prec<-1/(tau.sq)
beta ~ dnorm(0.0,1.0E-6)

d.uncent <- d - beta*(-3.7)
eqpoint <-d.uncent/(-beta)

d.new ~dnorm(d, prec)
OR.new <-exp(d.new)
}
```

Al di là degli IPD possiamo sicuramente valutare l'impatto del trattamento rispetto alla gravità della malattia utilizzando il "baseline risk".

Quanto un trattamento è dannoso?

Sottraiamo all'effectsize lo slope*predittore e lo dividiamo per il negato dello slope e otteniamo il **baseline risk "dannoso"** (**eqpoint**)

La mediana di eqpoint è: -4.45

$$e(-4.45) / (1+e(-4.45)) = 0.01$$

Si ha meno dell'1% di probabilità che il nuovo trattamento risulti essere dannoso.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
eqpoint	-8.124	441.2	3.55	-17.36	-4.45	4.135	30001	15000

Smoking Cessation A vs B vs C vs D

Mixed Treatment Comparison	Studio	No Contact	Self-help	Consulenza Individuale	Consulenza di Gruppo
			pA	pB	pD
AB(3)	1	79/702	77/694		
	2	18/671	21/535		
	3	8/116	19/146		
AC(15)	4	2/106		363/714	
	5	58/549		9/205	
	..				
	17	69/1117		54/888	
ACD(1)	18	9/140		23/140	10/138
AD(1)	19	0/20			9/20
BC(1)	20		20/49	16/43	
BCD(1)	21		11/78	12/85	29/170
BD(1)	22		7/66		32/127
CD(2)	23			12/76	20/74
	24			9/55	2/26

Stima su Reti e Regressione Logistica

Smoking Cessation

Per i 2 studi sui trattamenti C e D stimiamo una correlazione diretta:

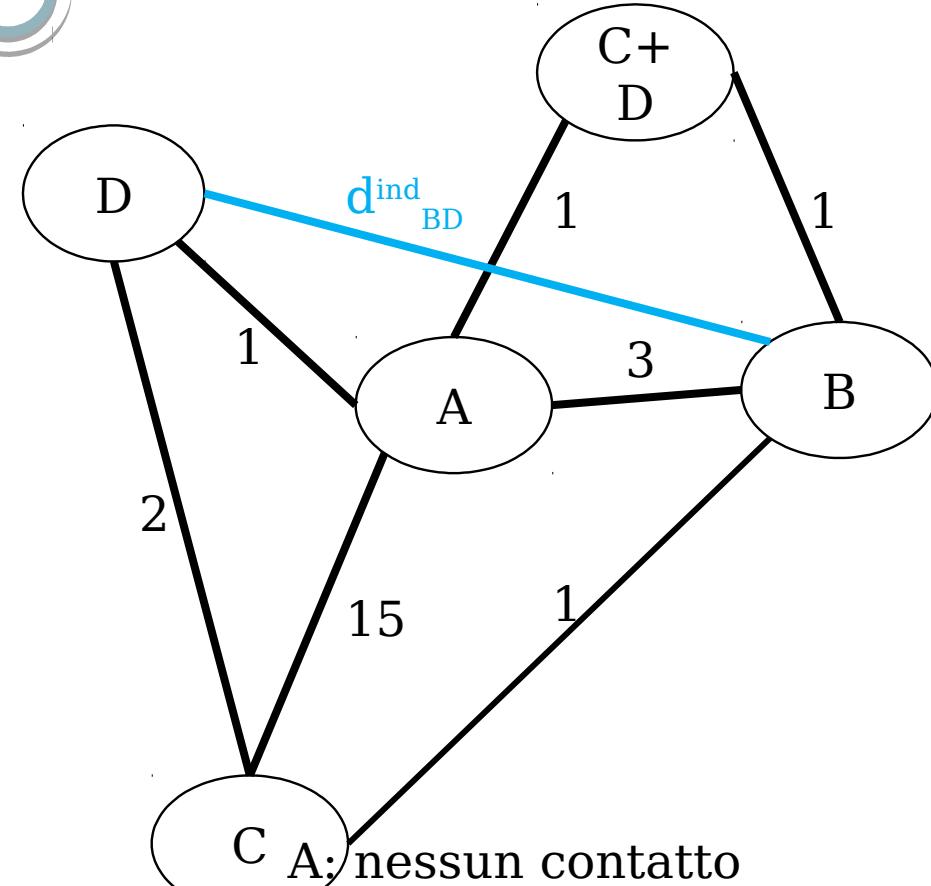
$$d^{\text{dir}}_{AB}$$

Tra B e D possiamo stimare una **correlazione indiretta** con le equazioni di consistenza

$$d^{\text{ind}}_{BD} = d^{\text{dir}}_{AD} - d^{\text{dir}}_{AB}$$

$$d^{\text{ind}}_{BC} = d^{\text{dir}}_{AC} - d^{\text{dir}}_{AB}$$

$$d^{\text{ind}}_{CD} = d^{\text{dir}}_{AD} - d^{\text{dir}}_{AC}$$



A: nessun contatto

B: aiuto autonomo

C: consulenza individuale

D: consulenza di gruppo

Esempio dello Smoking Cessation:



Fixed Effect Model

$$r_{jk} \sim \text{Binomial}(p_{jk}, n_{jk})$$

$$\text{logit}(p_{jk}) = \mu_j + d_{XY} I(k = Y)$$

$$d_{BC} = d_{AC} - d_{AB}$$

$$d_{BD} = d_{AD} - d_{AB}$$

$$d_{CD} = d_{AD} - d_{AC}$$

$$\mu_j, d_{AB}, d_{AC}, d_{AD} \sim \text{Normal}(0, 100^2)$$

Random Effect Model

$$r_{jk} \sim \text{Binomial}(p_{jk}, n_{jk})$$

$$\text{logit}(p_{jk}) = \mu_j + \delta_{jXY} I(k = Y)$$

$$\delta_{jXY} \sim \text{Normal}(\text{d}_{XY}, \sigma^2)$$

$$\sigma^2 \sim \text{Uniform}(0, 2)$$

$$d_{BC} = d_{AC} - d_{AB}$$

$$d_{BD} = d_{AD} - d_{AB}$$

$$d_{CD} = d_{AD} - d_{AC}$$

$$\mu_j, d_{AB}, d_{AC}, d_{AD} \sim \text{Normal}(0, 100^2)$$

Esempio dello Smoking Cessation:

s[] studio

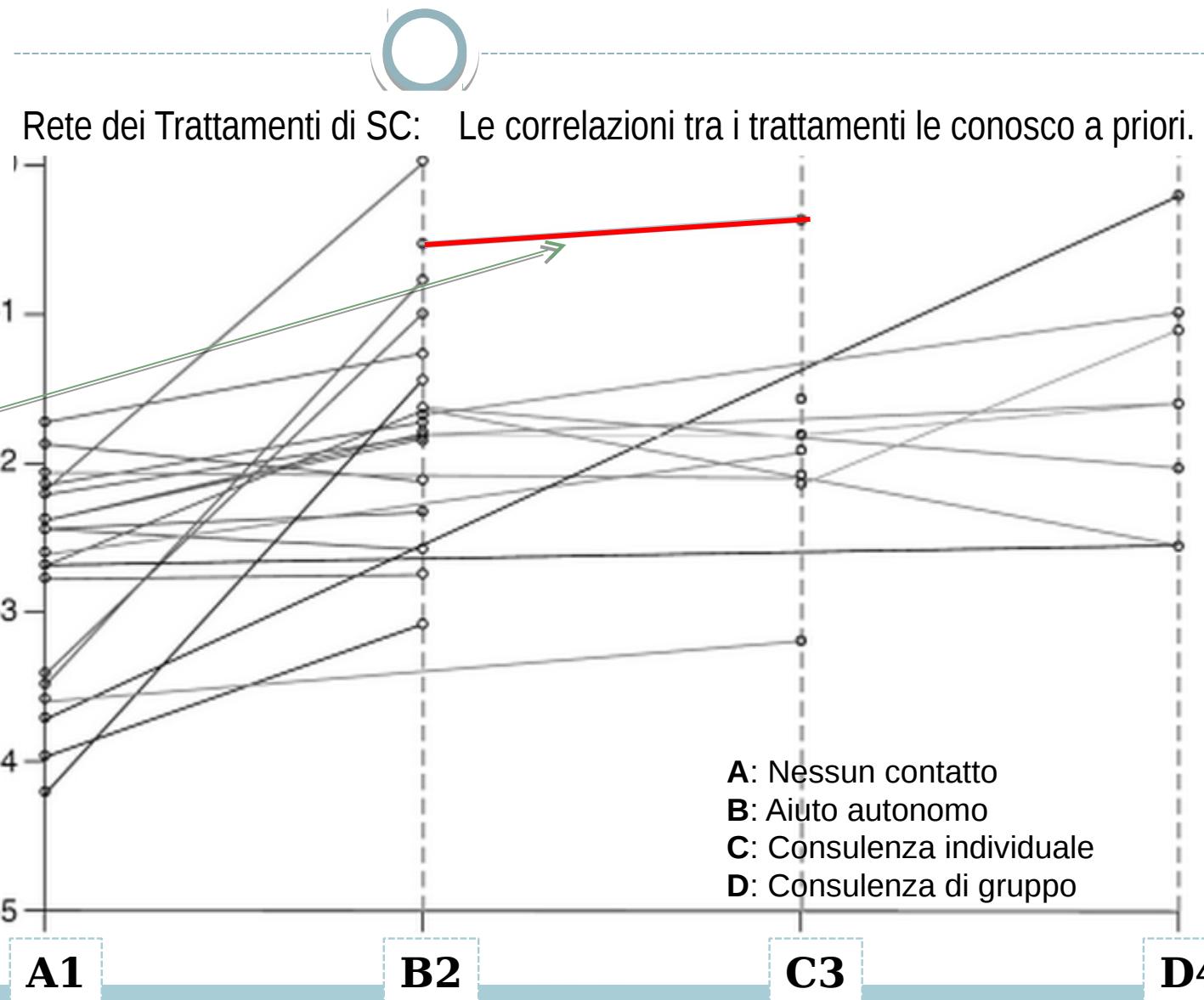
t[] trattamento

r[] numeratore

n[] denominatore

b[] trattamento correlato

s[]	t[]	r[]	n[]	b[]
1	1	79	702	1
1	2	77	694	1
2	1	18	671	2
2	3	16	43	2
2	4	29	170	2
3	1	75	731	1
3	3	363	714	1
4	1	2	106	1
.....				
19	3	73	675	1
20	1	69	1177	1
20	3	54	888	1
21	2	20	49	2
21	3	16	43	2
22	2	7	66	2
22	4	32	127	2
23	3	12	76	3
23	4	20	74	3
24	3	9	55	3
24	4	3	26	3



Esempio dello Smoking Cessation:



Non presenteremo il codice del modello. Ma i risultati ci dicono che rispetto ad A in termini di efficacia relativa di B,C,D quello più efficiente è D-dT[1].

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
d[2]	0.2251	0.1261	0.001714	-0.02124	0.2254	0.4724	10001	20000
d[3]	0.7657	0.058	9.583E-4	0.6524	0.7654	0.881	10001	20000
d[4]	0.8428	0.1743	0.002265	0.5	0.8425	1.183	10001	20000

In termini di efficacia assoluta del trattamento introducendo qualche

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
T[1]	0.07313	0.02635	1.956E-4	0.03408	0.06893	0.135330001	20000	
T[2]	0.09035	0.03362	7.79E-4	0.04044	0.08485	0.17	30001	20000
T[3]	0.14410	0.04775	3.762E-4	0.06952	0.13740	0.254330001	20000	
T[4]	0.15470	0.05484	4.995E-4	0.07033	0.14670	0.283430001	20000	

Modelli Markoviani e MetaAnalisi per Studi Omogenei



Trattamento preventivo per l'asma:

STW: una settimana di trattamento soddisfacente

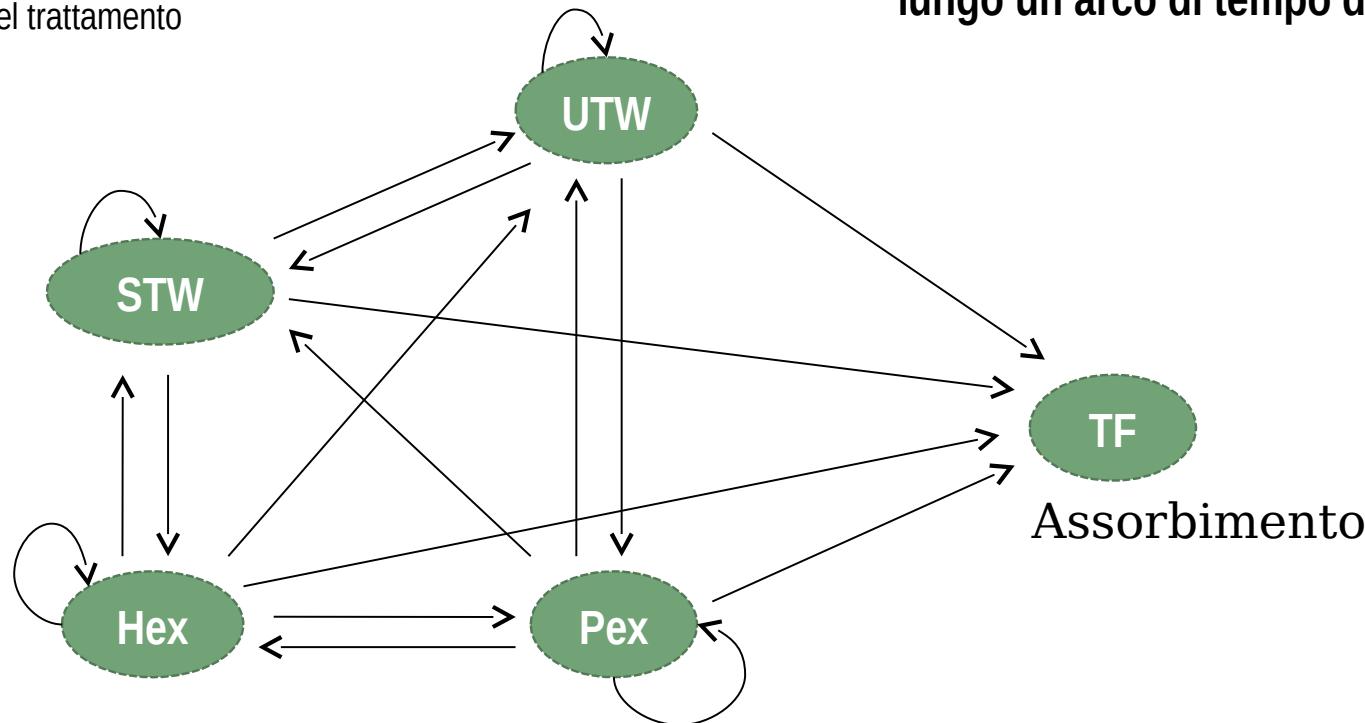
UTW: una settimana di trattamento insoddisfacente

Hex: peggioramento in ospedale in seguito al trattamento

Pex: peggioramento dopo cure primarie

TF: fallimento del trattamento

Vogliamo vedere se il trattamento con **Seretide™** è più efficace del trattamento con **Fluticasone** lungo un arco di tempo discreto **t**



Osservazioni sul numero di transizioni r_{ij}	STW $j=1$	UTW $j=2$	HEX $j=3$	Pex $j=4$	TF $J=5$	Totale per riga n_i
---	--------------	--------------	--------------	--------------	-------------	--------------------------

Trattamento col farmaco Seretide™ per 12 settimane

STW $i=1$	210	60	0	1	1	272
UTW $i=2$	88	641	0	4	13	746
HEX $i=3$	0	0	0	0	0	0
Pex $i=4$	1	0	0	0	1	2
TF $i=5$	0	0	0	0	81	81

Trattamento col farmaco Fluticasone per 12 settimane

STW $i=1$	66	32	0	0	2	100
UTW $i=2$	42	752	0	5	20	819
HEX $i=3$	0	0	0	0	0	0
Pex $i=4$	0	4	0	1	0	5
TF $i=5$	0	0	0	0	156	156

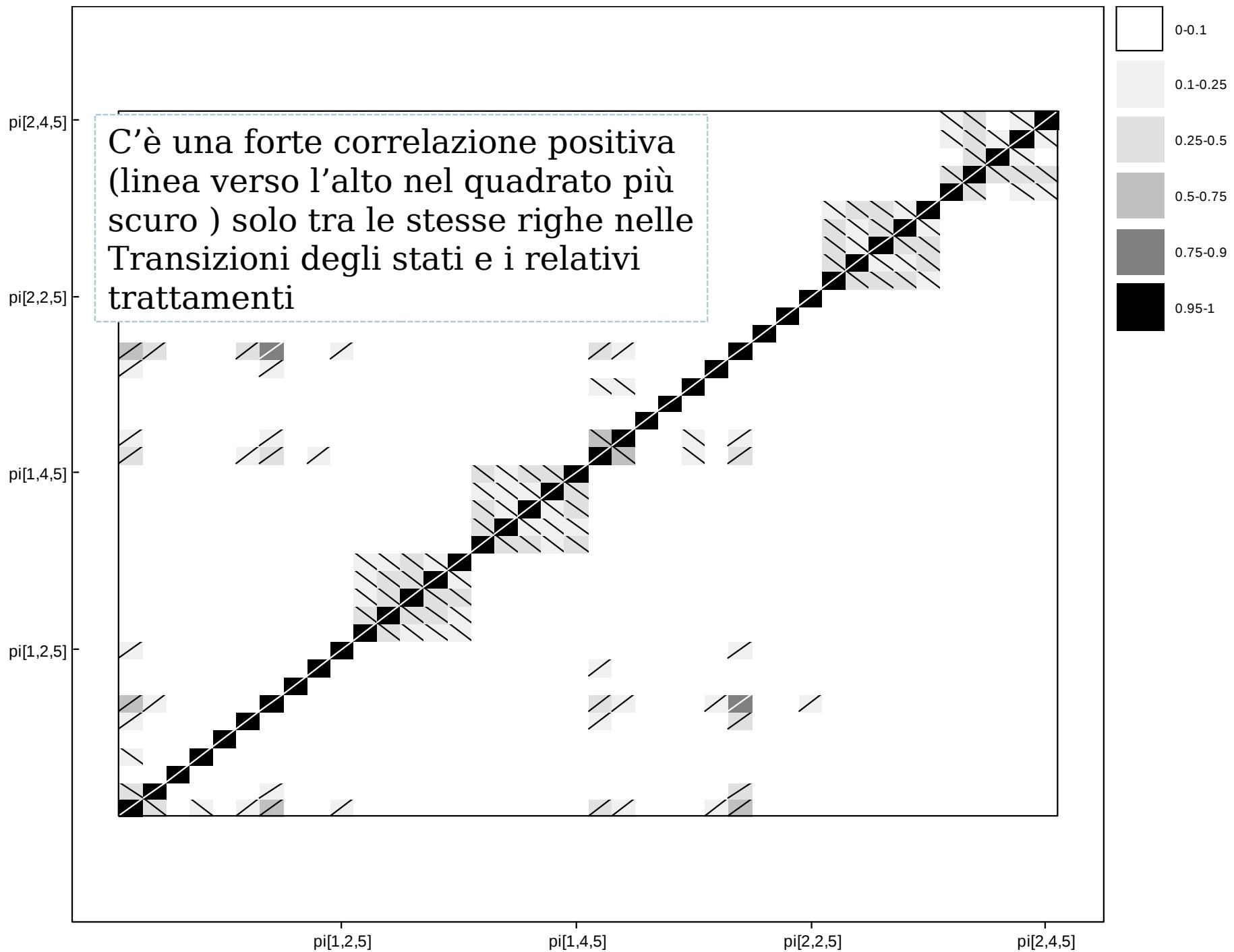
Ogni riga nella tabella è indipendente e le transizioni sono esclusive tra loro, per modellare ciò su più stati utilizziamo la multinomiale:

$$(r_{i,1}, r_{i,2}, r_{i,3}, r_{i,4}, r_{i,5}) \sim \text{Multinomial}(\Pi_{i,1}, \Pi_{i,2}, \Pi_{i,3}, \Pi_{i,4}, \Pi_{i,5} : n_i)$$

la distribuzione a priori coniugata è:

$$(\Pi_{i,1}, \Pi_{i,2}, \Pi_{i,3}, \Pi_{i,4}, \Pi_{i,5}) \sim \text{Dirichlet}(d_{i,1}, d_{i,2}, d_{i,3}, d_{i,4}, d_{i,5})$$





Risultati per le $\Pi_{i,1}$, $\Pi_{i,2}$, $\Pi_{i,3}$, $\Pi_{i,4}$, $\Pi_{i,5}$



Table 10.4 Posterior means for the transition probabilities with 95% credible intervals for treatment arms Seretide™ and Fluticasone in the asthma example.

Posterior mean for $\pi_{i,j}$ (95% credible interval)	STW, $j = 1$	UTW, $j = 2$	Hex, $j = 3$	Pex, $j = 4$	TF, $j = 5$
<i>Seretide™</i>					
STW, $i = 1$	0.762 (0.71,0.81)	0.220 (0.17,0.27)	0.004 (0.00,0.01)	0.007 (0.00,0.02)	0.007 (0.00,0.02)
UTW, $i = 2$	0.119 (0.10,0.14)	0.855 (0.83,0.88)	0.001 (0.00,0.01)	0.007 (0.00,0.01)	0.019 (0.01,0.03)
Hex, $i = 3$	0.200 (0.01,0.61)	0.200 (0.01,0.61)	0.200 (0.01,0.61)	0.200 (0.01,0.61)	0.200 (0.01,0.61)
Pex, $i = 4$	0.286 (0.04,0.64)	0.144 (0.00,0.47)	0.144 (0.00,0.46)	0.142 (0.00,0.46)	0.284 (0.04,0.64)
TF, $i = 5$	0	0	0	0	1
<i>Fluticasone</i>					
STW, $i = 1$	0.638 (0.54,0.73)	0.314 (0.23,0.41)	0.010 (0.00,0.03)	0.009 (0.00,0.03)	0.029 (0.01,0.07)
UTW, $i = 2$	0.052 (0.04,0.07)	0.914 (0.89,0.93)	0.001 (0.00,0.00)	0.007 (0.00,0.01)	0.026 (0.02,0.04)
Hex, $i = 3$	0.200 (0.01,0.60)	0.200 (0.01,0.60)	0.200 (0.01,0.60)	0.200 (0.01,0.60)	0.200 (0.01,0.60)
Pex, $i = 4$	0.101 (0.00,0.34)	0.500 (0.21,0.79)	0.100 (0.00,0.34)	0.199 (0.03,0.48)	0.100 (0.00,0.33)
TF, $i = 5$	0	0	0	0	1

Come utilizziamo le $\Pi_{i,1}, \Pi_{i,2}, \Pi_{i,3}, \Pi_{i,4}$, $\Pi_{i,5}$?



Esempio dopo 12 settimane come si evolve il modello? Qual è la proporzione di individui che si trova in un determinato stato?

Impostiamo $s_{i,t}$ come la proporzione probabilistica che un numero di individui si trova nello stato i al tempo t.

Tutti gli individui si troveranno in uno stato iniziale $t=1$: $s_{1,1} = 1$ e $s_{2,1} = s_{3,1} = s_{4,1} = 0$

al tempo $t=t+1$: $s_{1,t} = s_{1,(t-1)}\Pi_{1,i} + s_{2,(t-1)}\Pi_{2,i} + s_{3,(t-1)}\Pi_{3,i} + s_{4,(t-1)}\Pi_{4,i}$ per ogni i-stato.

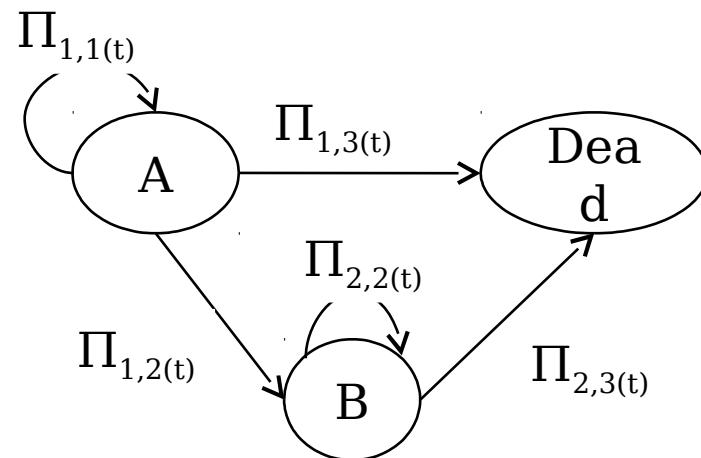
Il valore atteso che mi dice il numero di settimane nelle quali si è rimasti in un determinato stato :

$$E_i = \sum_{t=2}^{13} s_{i,t}$$

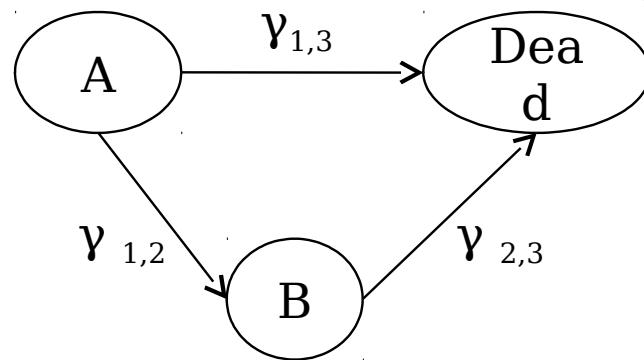
Da un semplice confronto è stato il migliore:

$$D_i = E_{tmt=1,i} - E_{tmt=2,i}$$

Modelli Makoviani e MetaAnalisi per Studi Eterogenei



5 probabilità



3 frequenze

Quello che potrebbe essere in comune tra i diversi studi eterogenei tra loro sono soltanto gli **stati** di una malattia. Per questa ragione se volessimo modellare i parametri di transizione rispetto a delle probabilità rischieremo di denaturare il modello. La soluzione sta nel modellare le frequenze di transizione.

Impostiamo il modello in modo tale che:

- 1) Per prima cosa ottengo i Parametri Basici dalla riparametrizzazione delle **Frequenze di Transizione γ**

E dopo ottengo i Parametri Funzionali che sono le **Probabilità di Transizione Π** e sono ricavate dalle γ tramite le equazioni di Kolmogorov

Tra loro parametri basici e parametri funzionali sono correlati tramite sistemi di equazioni.

$$A \xrightarrow{\gamma_{1,3}} \text{Dead}$$

$$A \xrightarrow{\gamma_{1,2}} B$$

$$\frac{\gamma_{1,2}}{\gamma_{1,3}} +$$

$$B \xrightarrow{\gamma_{2,3}} \text{Dead}$$

$$B \xrightarrow{\gamma_{2,3}}$$

2 frequenze

Relazioni tra γ e Π



La relazione che intercorre tra **Frequenze di Transizione γ** e **Probabilità di Transizione Π** si ritrova nelle equazioni di Kolmogorov in avanti (riescono a descrivere i processi stocastici dinamici)

Sono equazioni differenziali

$$\mathbf{P}^I(t) = \mathbf{P}(t)\mathbf{G}$$

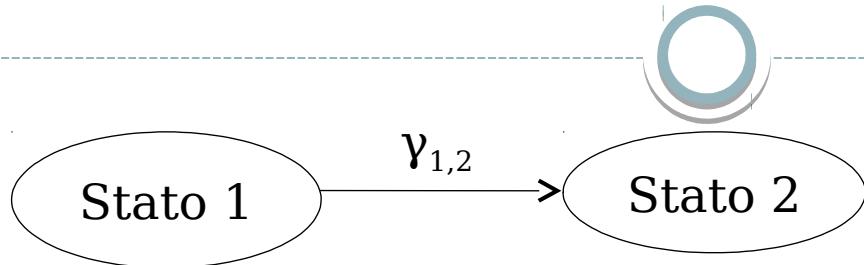
$\mathbf{P}(t)$ è la **matrice di Probabilità di Transizione** in un ciclo di lunghezza t , **\mathbf{G}** è una matrice delle **Frequenze di Transizione**.

La matrice di transizione **$\mathbf{P}(t)$** può essere definita come l'esponente di t volte la matrice **\mathbf{G}** :

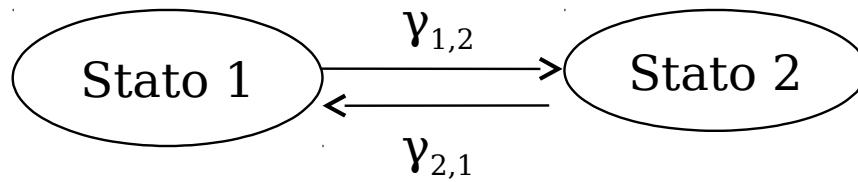
$$\mathbf{P}(t) = e^{t\mathbf{G}}$$

Per alcuni modelli comuni si può calcolare.

Relazioni tra γ e Π



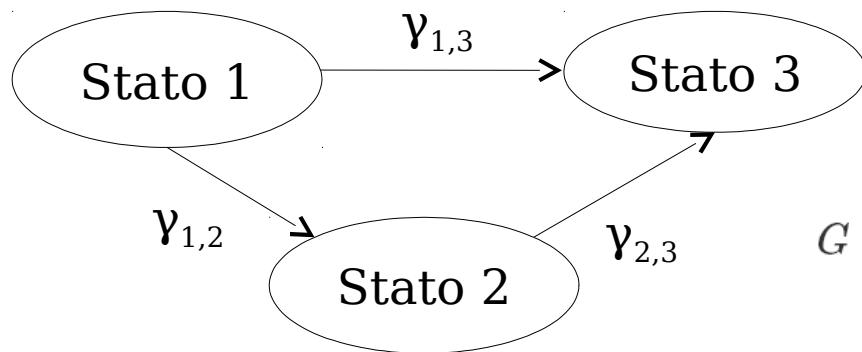
$$P(t) = \begin{bmatrix} e^{-(\gamma_{1,2})t} & 1 - e^{-(\gamma_{1,2})t} \\ 0 & 1 \end{bmatrix}$$



$$G = \begin{bmatrix} -\gamma_{1,2} & \gamma_{1,2} \\ \gamma_{2,1} & -\gamma_{2,1} \end{bmatrix}$$

$$P(t) = \begin{bmatrix} \frac{\gamma_{2,1} + \gamma_{1,2}e^{-(\gamma_{1,2}+\gamma_{2,1})t}}{\gamma_{2,1} + \gamma_{1,2}} & 1 - \frac{\pi_{1,1}(t)}{\gamma_{1,2} + \gamma_{2,1}e^{-(\gamma_{1,2}+\gamma_{2,1})t}} \\ 1 - \frac{\pi_{2,2}(t)}{\gamma_{2,1} + \gamma_{1,2}} & \frac{\gamma_{1,2} + \gamma_{2,1}e^{-(\gamma_{1,2}+\gamma_{2,1})t}}{\gamma_{2,1} + \gamma_{1,2}} \end{bmatrix}$$

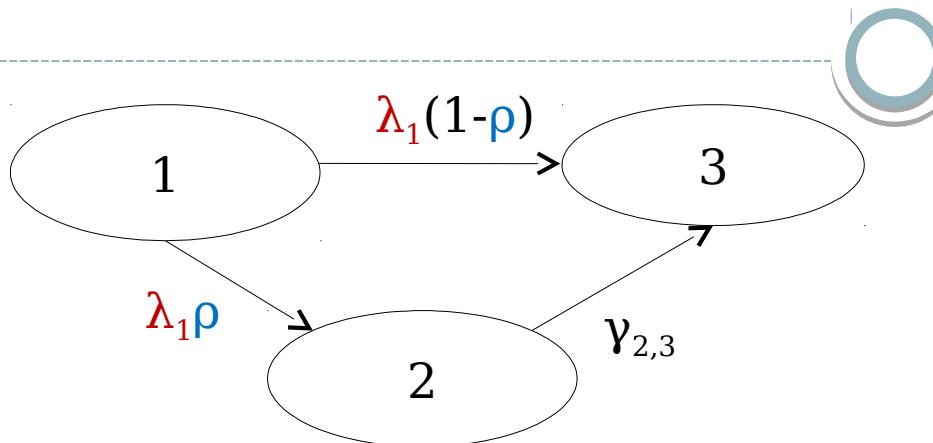
Relazioni tra γ e Π



$$G = \begin{bmatrix} -\gamma_{1,2} + \gamma_{1,3} & \gamma_{1,2} & \gamma_{1,3} \\ 0 & -\gamma_{2,3} & \gamma_{2,3} \\ 0 & 0 & 0 \end{bmatrix}$$

$$P(t) = \begin{bmatrix} e^{-\lambda_1 t} & \frac{\gamma_{1,2} e^{-(\gamma_{2,3})t} (1 - e^{-(\lambda_1 - \gamma_{2,3})t})}{\gamma_{1,2} + \gamma_{1,3} - \gamma_{2,3}} & (1 - \pi_{1,1}(t) - \pi_{1,2}(t)) \\ 0 & e^{-(\gamma_{2,3})t} & (1 - e^{-(\gamma_{2,3})t}) \\ 0 & 0 & 1 \end{bmatrix}$$

Esempio Con Modello a 3 Stati



Riparametrizzazione di
 $\gamma_{1,3} = \lambda_1(1-\rho)$ e $\gamma_{1,2} = \lambda\rho$

λ_1 è variabile tra gli studi.

Si assumono essere variabili fisse tra gli studi

$$s_i \quad \text{Log}(\lambda_1) \sim \text{Normal}(L, \tau^2)$$

$$\rho \sim \text{Beta}(1,1)$$

$$\gamma_{2,3} \sim \text{Exponential}(.001)$$

A priori

$$L \sim \text{Normal}(0, 100^2)$$

$$\tau \sim \text{Uniform}(0,1)$$

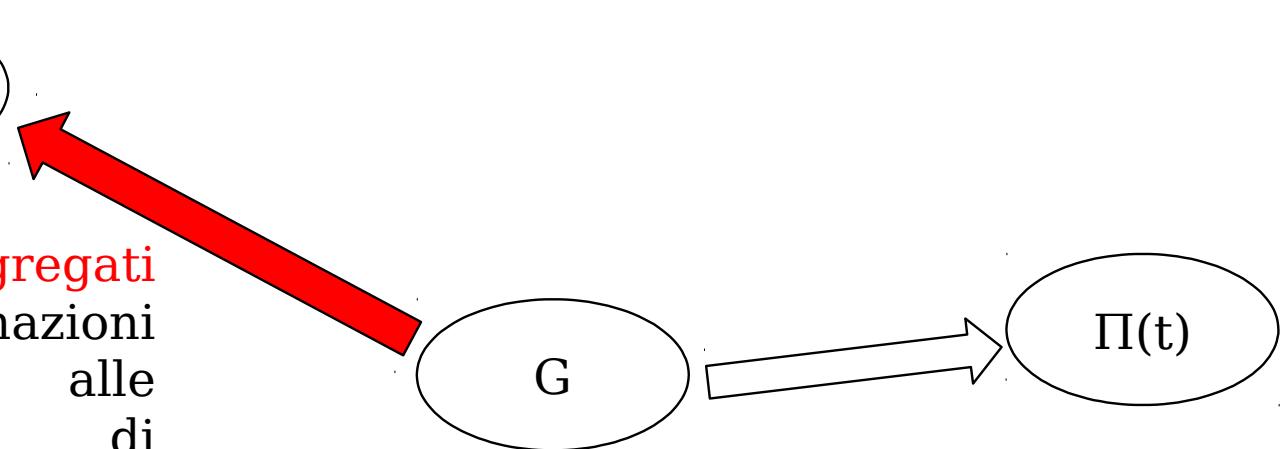
Random Effect Model
Welton & Ades

Abbiamo tre parametri basici: λ_1 , ρ e $\gamma_{2,3}$ tramite questi possiamo ricavare gli altri parametri funzionali sia di frequenza che transizione.

Esempio Con Modello a 3 Stati



I **dati aggregati** forniscono informazioni direttamente alle probabilità di transizione, che sono funzioni dei parametri di G .
Sono osservazioni fatte per dei cicli di lunghezza t_x .



I **dati storici** forniscono informazioni direttamente ai parametri di G . Ovviamente sia se sono dati sulla prima transizione sia per transizioni successive.

Esempio Con Modello a 3 Stati



Studi che riportano le transizioni. T=1		Stato di destinazione			Questi studi riportano solo transizioni che portano alla morte nello stato 3.			
		j = 1	j = 2	j = 3	Studio s	Numero transizioni sullo stato 3, $r_s(T)$	Numero di individui che partono dallo stato 1, $n_s(T)$	# Osservazioni sulla lunghezza ciclo, T
1	i = 1	87	4	9	3	38	181	2
	i = 2	0	3	1	4	15	177	1.5
2	i = 1	210	8	15	5	11	103	2
	i = 2	0	5	3	6	13	335	5

due studi ognuno dei quali riporta il numero di transizioni per il ciclo T=1 (una settimana)

$(r_{s,i,1}, r_{s,i,2}, r_{s,i,3}) \sim \text{Multinomial}(\Pi_{s,i,1}(T), \Pi_{s,i,2}(T), \Pi_{s,i,3}(T); n_{s,i})$
 Π tra loro sono probabilità competitive,

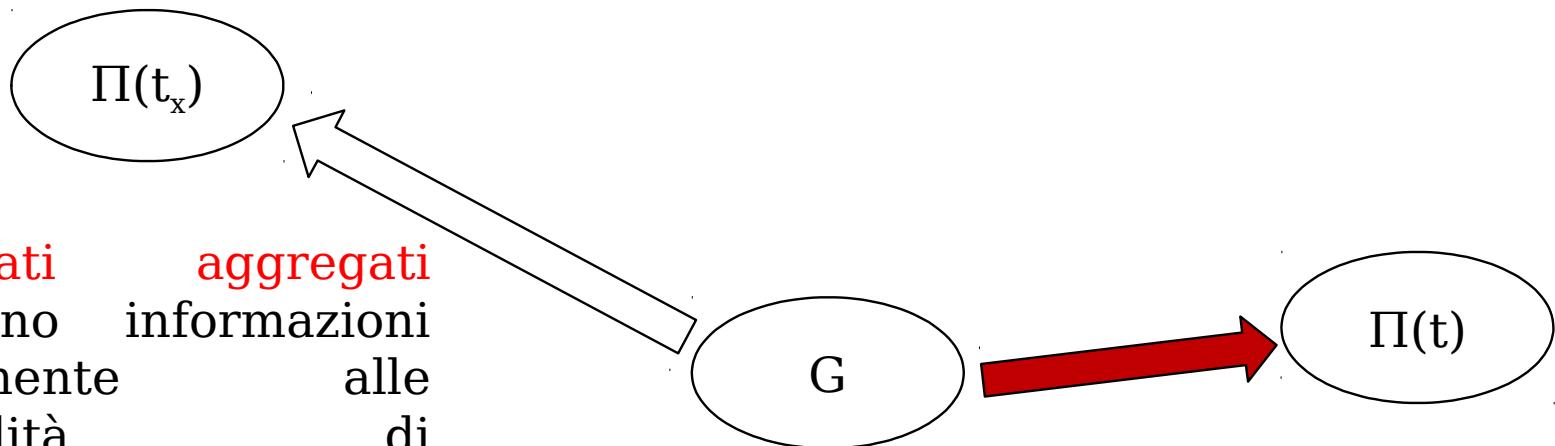
$n_{s,i}$ numero totale per riga di individui che lasciano quello stato per settimana.

Gli associamo

$$r_s(T) \sim \text{Binomial}(\Pi_{s,1,3}(T); n_s(T))$$

Esempi di Dati Aggregati

Esempio Con Modello a 3 Stati



I **dati aggregati** forniscono informazioni direttamente alle probabilità di transizione, che sono funzioni dei parametri di G .
Sono osservazioni fatte per dei cicli di lunghezza t_x .

I **dati storici** forniscono informazioni direttamente ai parametri di G .
Ovviamente sia se sono dati sulla prima transizione sia per transizioni successive.

Esempi di Dati Storici

Studio s	Numero transizioni osservate $r_{1,2}$ $r_{1,3}$		Persone Annualmet e a rischio, E:
6	8		120
7	20		620
8		12	140
9		44	577
10		9	34
11		12	35
12		4	15
13		6	25

$$r_{s,1,2} \sim \text{Poisson}(\mu_s)$$

$$\log(\mu_s) = \log(E_s) + \log(\lambda_{1,s}(1-\rho))$$

$$r_{s,1,3} \sim \text{Poisson}(\mu_s)$$

$$\log(\mu_s) = \log(E_s) + \log(\lambda_{1,s}\rho)$$

$$r_{s,2,3} \sim \text{Poisson}(\mu_s)$$

$$\log(\mu_s) = \log(E_s) + \log(\gamma_{2,3})$$

Gli individui che lasciano lo stato 1

$$n_s = m_{s,1,2} + m_{s,1,3}$$

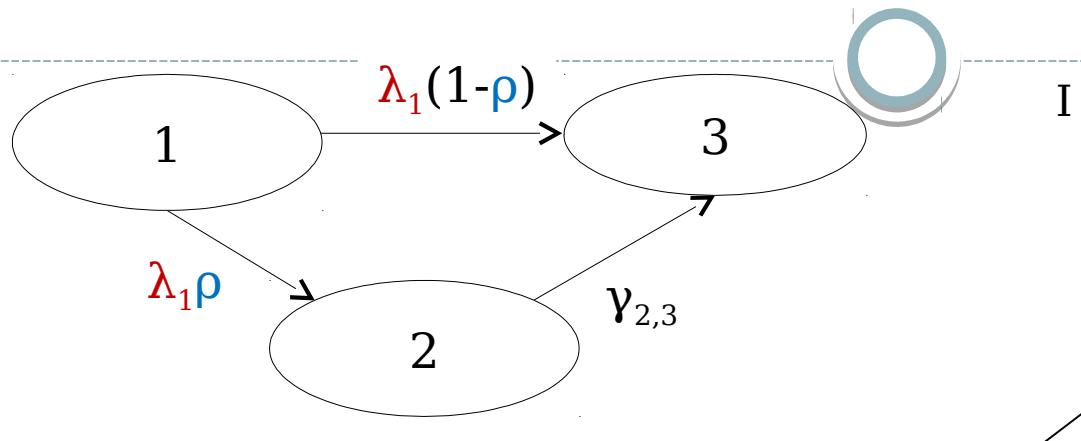
$$n_s \sim \text{Poisson}(\nu_s) \text{ con } \log(\nu_s) = \log(E_s) + \log(\lambda_{1,s})$$

In definitiva: $r_{s,1,3} \sim \text{Binomial}(\rho, n_s)$

Che contempla anche quella di non andare da 1 a 2

Studio s	Numero transizioni osservate $r_{1,2}$ $r_{1,3}$		Persone Annualmet e a rischio, E:
14	18	40	120
15	30	75	620

Esempio Con Modello a 3 Stati



Si assumono essere variabili fisse tra gli studi.

λ_1 è variabile tra gli studi.

$$s_i \quad \text{Log}(\lambda_1) \sim \text{Normal}(L, \tau^2)$$

$$\rho \sim \text{Beta}(1,1)$$

$$\gamma_{2,3} \sim \text{Exponential}(.001)$$

A priori

$$L \sim \text{Normal}(0, 100^2)$$

$$\tau \sim \text{Uniform}(0,1)$$

I risultati alternativi del modello:

$$\gamma_{1,2} = e^L (1-\rho)$$

$$\gamma_{1,3} = e^L \rho$$

$$\gamma_{1,2} = e^{\text{pred}} (1-\rho)$$

$$\gamma_{1,3} = e^{\text{pred}} \rho$$

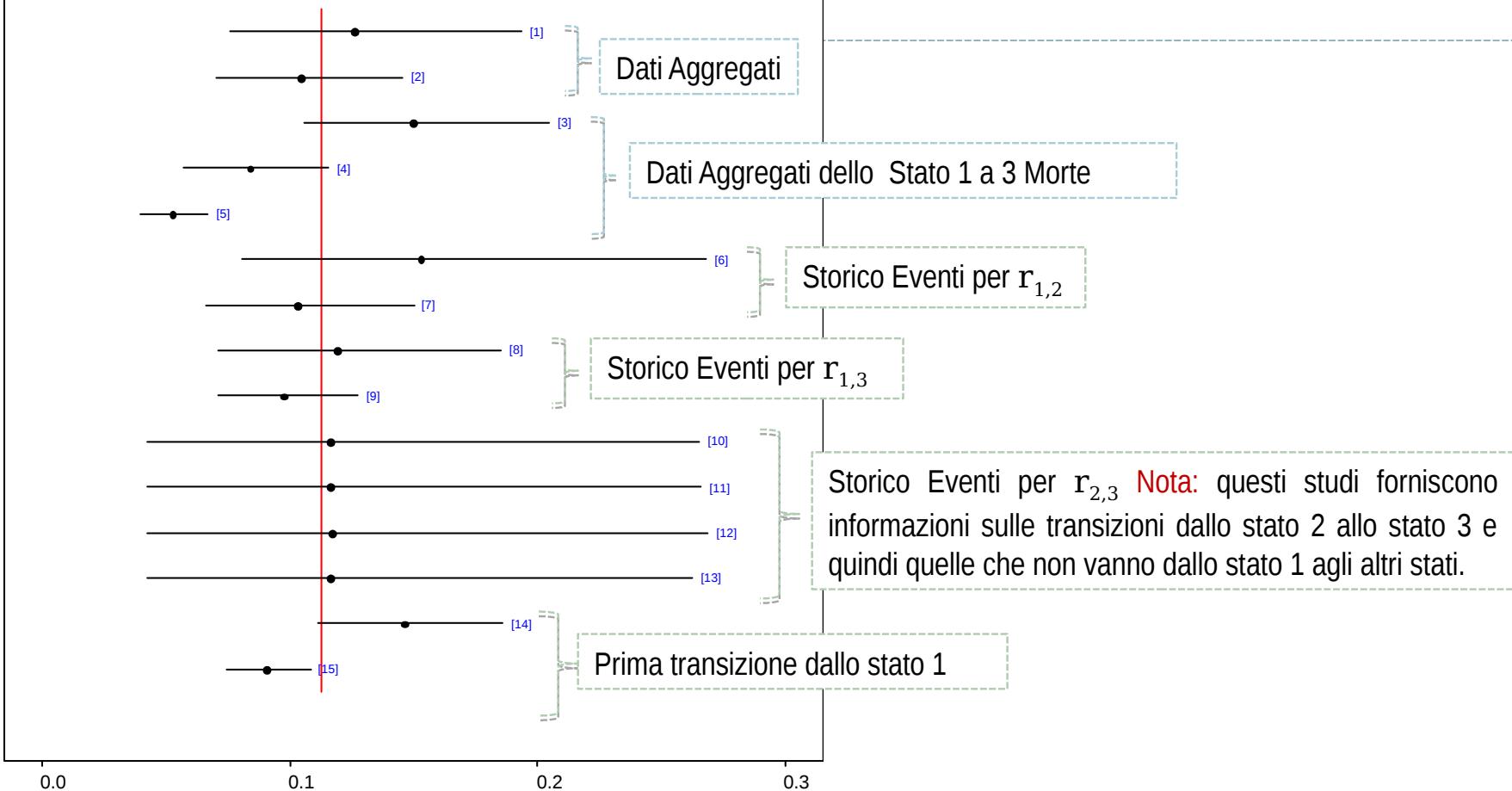
Dove $\text{Log}(\lambda_1^{\text{pred}}) \sim \text{Normal}(L, \tau^2)$

$$\gamma_{1,2} = \lambda_{1,15} (1-\rho)$$

$$\gamma_{1,3} = \lambda_{1,15} \rho$$

Random Effect Model
Welton & Ades

caterpillar plot: lambda1



Impostato il modello e facendo diverse assunzioni che non vedremo ottengo una stima dei parametri Basici

	L	τ	ρ	$\gamma_{2,3}$
95% CrI	-2.25 (-2.54, -1.96)	0.41 (0.22, 0.73)	0.68 (0.62, 0.74)	0.30 (0.21, 0.41)

Ricavo le restanti frequenze di transizione $\gamma_{1,3}$ e $\gamma_{1,2}$:

Media a posteriori 95% CrI	$\gamma_{1,2}$	$\gamma_{1,3}$
Media Casuale L	0.034 (0.024, 0.047)	0.072 (0.053, 0.097)
Distribuzione Predittiva	0.037 (0.013, 0.086)	0.079 (0.028, 0.183)
Calibrazione rispetto lo studio 15 (che è quello più recente)	0.029 (0.022, 0.037)	0.062 (0.050, 0.075)

Applico le equazioni in avanti di Kolmogorov con le **frequenze di transizione** ottenute per un ciclo di lunghezza T esempio per T= 1 e ottengo una stima delle **probabilità di transizione**:

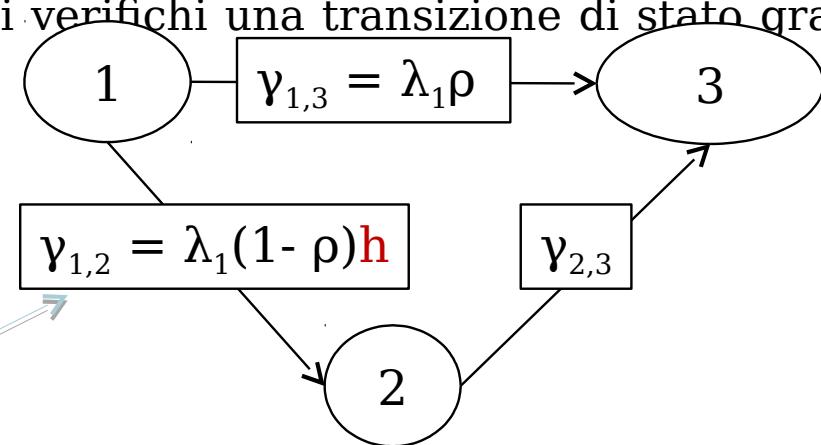
Media a posteriori 95% CrI	$\Pi_{1,1}$	$\Pi_{1,2}$	$\Pi_{1,3}$	$\Pi_{2,2}$	$\Pi_{2,3}$
Media Casuale L	0.90 (0.87, 0.92)	0.03 (0.03, 0.04)	0.07 (0.05, 0.10)	0.74 (0.05, 0.10)	0.26 (0.19, 0.34)
Distribuzione Predittiva	0.89 (0.77, 0.96)	0.03 (0.01, 0.06)	0.08 (0.03, 0.17)		
Calibrazione rispetto lo studio 15	0.91 (0.90, 0.93)	0.02 (0.02, 0.03)	0.06 (0.05, 0.08)		

L'Impatto del Trattamento sulla Transizione

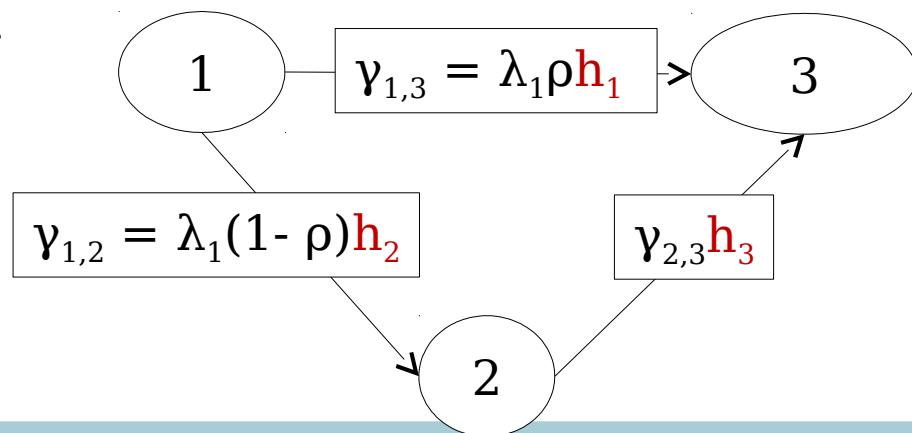


- L'effetto del trattamento può influenzare una singola transizione di stato , due , tutte...
- **h** = Introduciamo h come il rischio che si verifichi una transizione di stato grazie al trattamento rispetto al controllo.

Se nello studio di questo modello così strutturato ottengo ad esempio una media a posteriori per h di 0.525 95%CrI significa che stiamo riducendo di più della metà la frequenza che un Individuo transiti dallo stato 1 allo stato 2.



Oppure...



Modelli Gerarchici ed Empirici Bayesiani

Meta-Analisi tra diversi tipologie di studi osservazionali e RCT e poi ne ricaviamo un effect-size riassuntivo.

- Empirical Bayes (EB) Approccio bayesiano empirico
- Hierarchical Bayes (HB) Approccio bayesiano gerarchico.

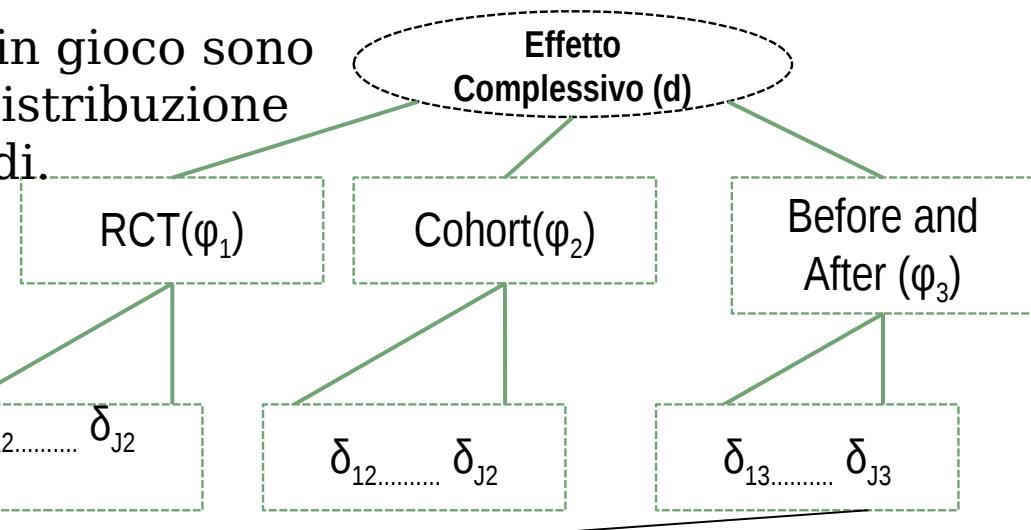
In EB tutti i parametri θ seguono una legge nota a priori a meno di un parametro λ giacchè $p(\theta | \lambda)$

In HB invece tutte le quantità in gioco sono considerate aleatorie, la loro distribuzione viene specificata in diversi stadi.

- 1) $p(\underline{y}_j | \theta_j)$
- 2) $p(\theta_1, \dots, \theta_K | \lambda)$
- 3) $p(\lambda_1, \lambda_2) = p_1(\lambda_1 | \lambda_2)p_2(\lambda_2)$

Se gli iperparametri sono 2...

individuiamo anche qui causa-effetto



Comorbidity



Se due malattie i e j hanno più probabilità di manifestarsi in un individuo rispetto che una sola allora parliamo di comorbidità.

La distanza si può misurare in termini di rischio relativo dove C è il numero di pazienti affetti da entrambe le malattie, N è il numero totale di individui e le P sono le incidenze delle malattie.

$$RR_{ij} = \frac{C_{ij}N}{P_i P_j}$$

o come una Φ -correlation:

$$\phi_{ij} = \frac{C_{ij}N - P_i P_j}{\sqrt{P_i P_j (N - P_i)(N - P_j)}}$$

Il problema è che le Mappe di Comorbidity con indici così semplicistici non rispecchiano la complessità -omica e clinica. Degli indici più complessi potrebbero essere ricavati con meta-analysis attraverso Modelli Gerarchici

Comorbidity



Dove dalle informazioni non è chiaro individuare la direzionalità della progressione di una malattia si utilizzano gli OR_{ij} su popolazione α e β

$$OR_{ij}(\alpha, \beta) = \frac{p_{ij}(\alpha)(1-p_{ij}(\beta))}{p_{ij}(\beta)(1-p_{ij}(\alpha))},$$

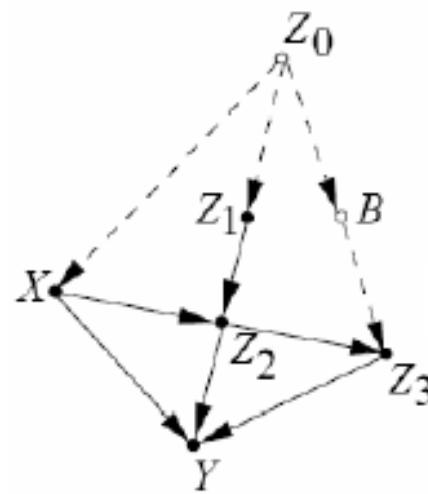
$$p_{ij}(\alpha) = C_{ij}(\alpha) / N_\alpha$$

è la probabilità dell'occorrenza delle malattie i
e j sulla popolazione α

Multi-Comorbidity e Qualche Ipotesi



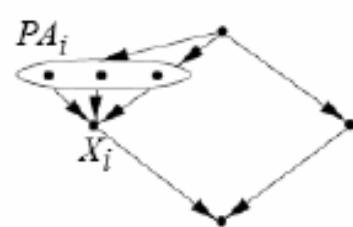
Ottenere degli indici di comorbidità che rispecchino la causalità (Reti Bayesiane e Inferenza su Reti Bayesiane)



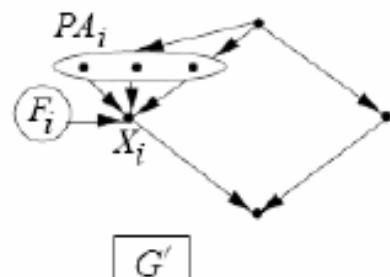
In questo semplice esempio vediamo il rapporto causa-effetto visto come una funzione che porta da X allo spazio delle probabilità di Y questa sarebbe : $P(y|\text{do}(x))$

$$P(y|\text{do}(x)) = \sum_{z_1} \sum_{z_2} \sum_{z_3} P(y|z_2, z_3, x) P(z_2 | z_1, x) \times \sum_{x'} P(z_3 | z_1, z_2, x') P(z_1, x')$$

Nel calcolo di queste probabilità ci sono dei criteri per non andare troppo indietro nei rapporti causa-effetto.



G



G'

Figure 3.2: Representing external intervention F_i by an augmented network $G' = G \cup \{F_i \rightarrow X_i\}$.

$$P(x_i | pa_i) = \begin{cases} P(x_i | pa_i) & \text{if } F_i = \text{idle}, \\ 0 & \text{if } F_i = do(x'_i) \text{ and } x_i \neq x'_i, \\ 1 & \text{if } F_i = do(x'_i) \text{ and } x_i = x'_i \end{cases}$$

$$P(x_1, \dots, x_n | \hat{x}'_i) = P'(x_1, \dots, x_n | F_i = do(x'_i))$$

where P' is represented by G'

Meta Analysis vs Meta Synthesis

MetaAnalysis vs MetaSynthesis [1]:

Applichiamo meta-sintesi tramite uno stimatore che da stime unbiased basandosi sui selection diagrams che riportano le informazioni che ogni studio possiede nel suo ambiente target.

Così per questa stima:

$$P^*(y|do(x)) = \sum_z P^*(y|x, z)P^*(z)$$

Uso direttamente g, oppure in combinazione b e d.

I pallini pieni sono le info fornite da ogni studio, i vuoti il contrario, i quadrati sono a priori esterne.

WinBugs  R(pcalg inferire causalità e altri package per reti bayesiane)

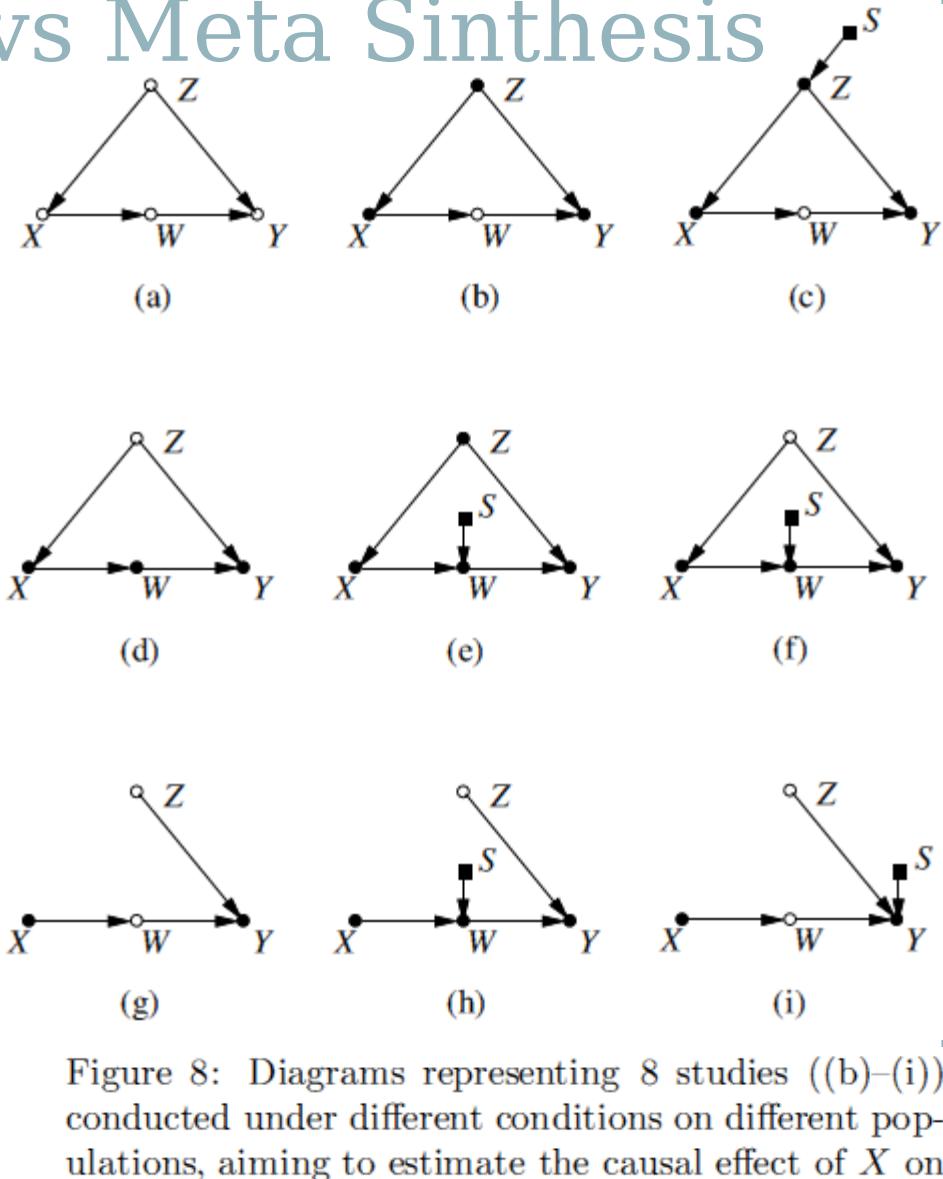


Figure 8: Diagrams representing 8 studies ((b)–(i)) conducted under different conditions on different populations, aiming to estimate the causal effect of X on Y in the target population, shown in 8(a).

MCMC and Gibbs Sampler e Probabilistic Graphical Models

- Campiono ciascuna VA del grafo, una alla volta.
- Esempio:
 - $x_1^{s+1} \sim p(x_1 | x_2^s, x_3^s)$
 - $x_2^{s+1} \sim p(x_2 | x_1^{s+1}, x_3^s)$
 - $x_3^{s+1} \sim p(x_3 | x_1^{s+1}, x_2^{s+1})$
- La $p(x_i | x_{-i})$ è la full conditional
- Questa viene semplificata usando la Markov Blanket

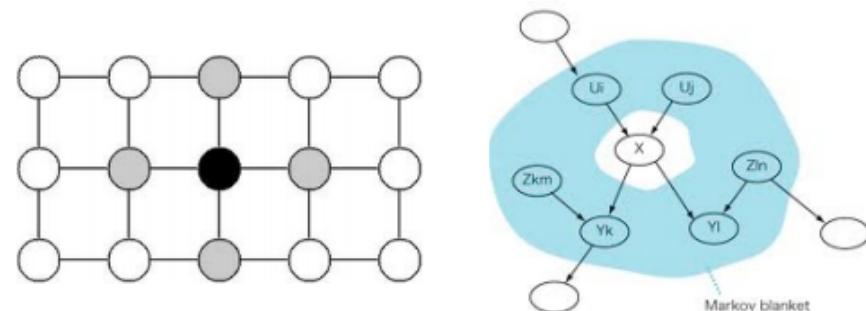
$$P(x_i | x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i | MB(x_i))$$

MB: Per un nodo A si prendono i figli, i parent(figli) e i parent(A).

```

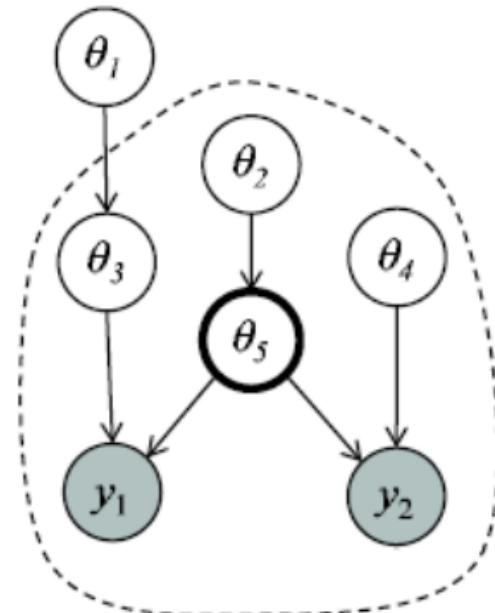
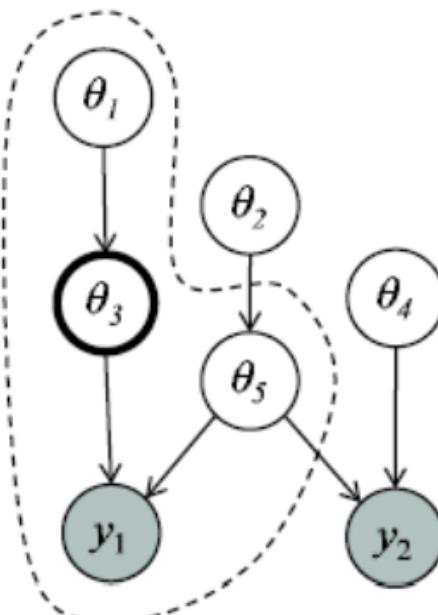
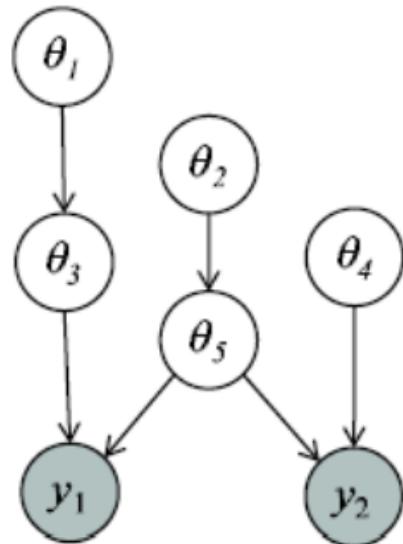
Procedure Gibbs-Sample (
   $X$  // Set of variables to be sampled
   $\Phi$  // Set of factors defining  $P_\Phi$ 
   $P^{(0)}(X)$ , // Initial state distribution
   $T$  // Number of time steps
)
1   Sample  $x^{(0)}$  from  $P^{(0)}(X)$ 
2   for  $t = 1, \dots, T$ 
3        $x^{(t)} \leftarrow x^{(t-1)}$ 
4   for each  $X_i \in X$ 
5       Sample  $x_i^{(t)}$  from  $P_\Phi(X_i | x_{-i})$ 
6       // Change  $X_i$  in  $x^{(t)}$ 
7   return  $x^{(0)}, \dots, x^{(T)}$ 

```



MCMC and Gibbs Sampler and PGM

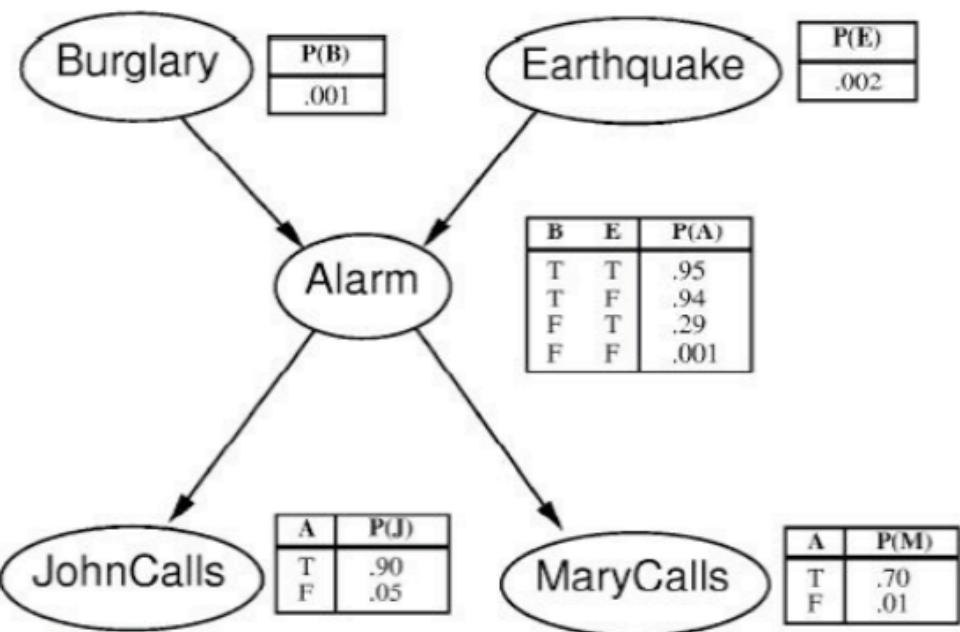
$$P[X_1, \dots, X_n] = \prod_{i=1}^n P[X_i | pa_i]$$



$$p(\theta_3 | \theta_1, \theta_2, \theta_4, \theta_5, y_1, y_2) = p(\theta_3, \theta_1, \theta_2, \theta_4, \theta_5, y_1, y_2)$$

$$p(\theta_3 | \theta_1, \theta_2, \theta_4, \theta_5, y_1, y_2) = p(\theta_1)p(\theta_2)p(\theta_4)p(\theta_3 | \theta_1)p(\theta_5 | \theta_2)p(y_1 | \theta_3, \theta_5)p(y_2 | \theta_5, \theta_4)$$

$$p(\theta_3 | \theta_1, \theta_2, \theta_4, \theta_5, y_1, y_2) = p(\theta_3 | \theta_1)p(y_1 | \theta_3, \theta_5)$$

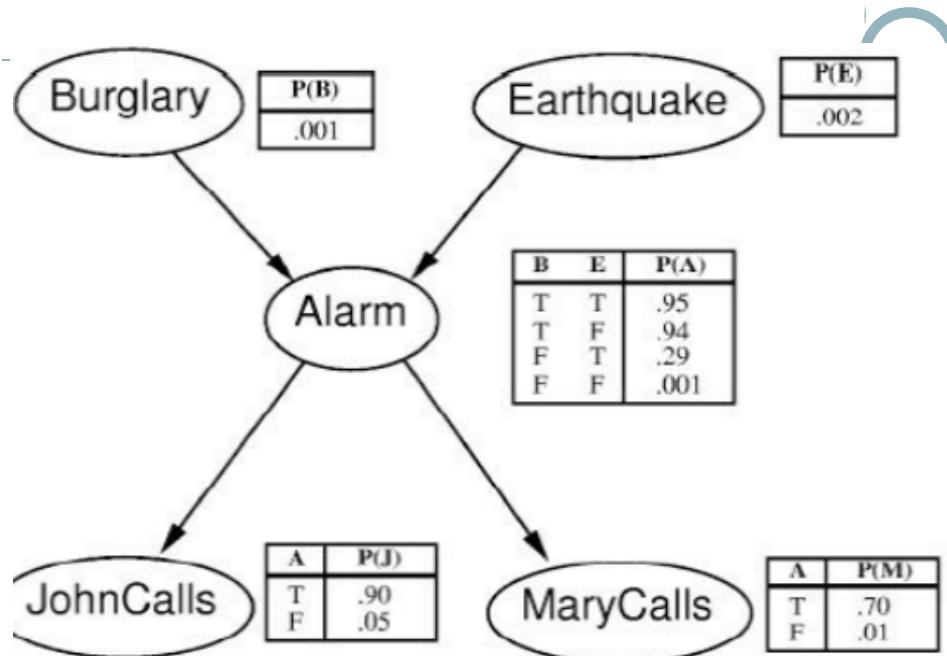


t	B	E	A	J	M
0	F	F	F	F	F
1					
2					
3					
4					

- Inizializzazione: ordine di visita B,E,A,J,M, t=0 tutte le variabili FALSE
 - t=1: campiono $P(B | E, A)$ con Bayes:
$$P(B | A, E) \propto P(A | B, E)P(B)$$
 - $A = \text{FALSE}$, $E = \text{FALSE}$
- $$P(B = T | A = F, E = F) \propto (0.06)(0.01) = 0.0006$$
- $$P(B = F | A = F, E = F) \propto (0.999)(0.999) = 0.9980$$

t	B	E	A	J	M
0	F	F	F	F	F
1		F			
2					
3					
4					

MCMC and Gibbs Sampler and PGM



t	B	E	A	J	M
0	F	F	F	F	F
1	F		T		
2					
3					
4					

- $t=1$: campiono $P(E | A, B)$ con Bayes: $P(E | A, B) \propto P(A | B, E)P(E)$
- $A = \text{FALSE}$, $B = \text{FALSE}$

$$P(E = T | A = F, B = F) \propto (0.71)(0.02) = 0.0142$$

$$P(E = F | A = F, B = F) \propto (0.999)(0.998) = 0.9970$$

MCMC and Gibbs Sampler and PGM

- t=1: campiono $P(A | B, E, J, M)$ con Bayes:

$$P(A | B, E, J, M) \propto P(J | A)P(M | A)P(A | B, E)$$

- $(B, E, J, M) = (F, T, F, F)$

$$\begin{aligned}P(A = T | B = F, E = T, J = F, M = F) \\ \propto (0.1)(0.3)(0.29) = 0.0087\end{aligned}$$

$$\begin{aligned}P(A = F | B = F, E = T, J = F, M = F) \\ \propto (0.95)(0.99)(0.71) = 0.6678\end{aligned}$$

- t=1: campiono $P(J | A)$ direttamente dalla CPT:

- $A = F$

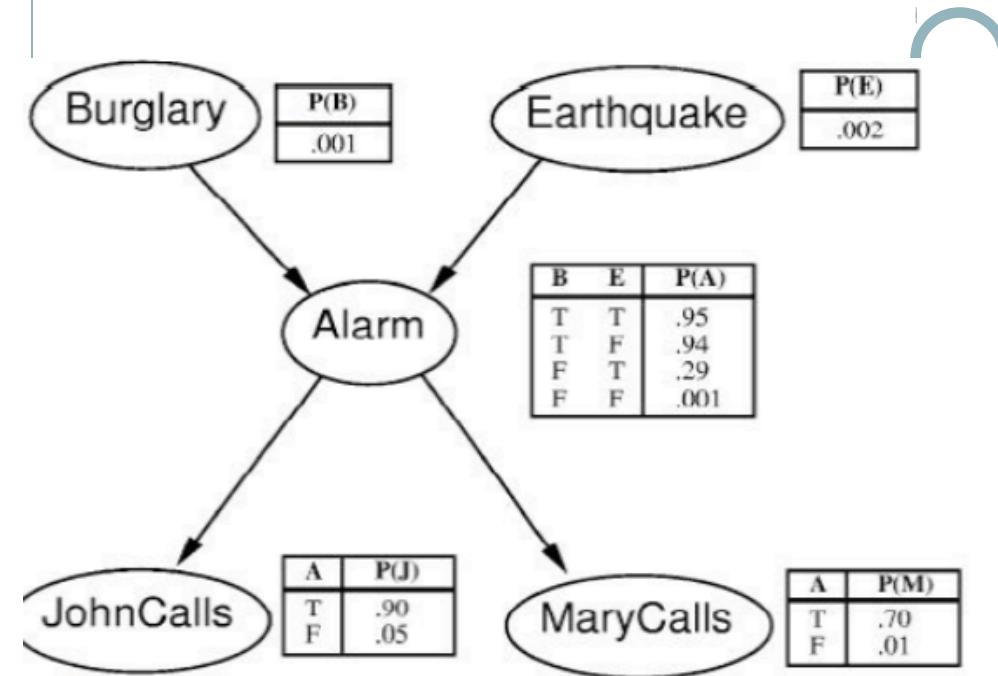
$$P(J = T | A = F) \propto 0.05$$

$$P(J = F | A = F) \propto 0.95$$

t	B	E	A	J	M
0	F	F	F	F	F
1	F	T	F		
2					
3					
4					

t	B	E	A	J	M
0	F	F	F	F	F
1	F	T	F	T	
2					
3					
4					

MCMC and Gibbs Sampler and PGM



A sampling table showing the state of variables B, E, A, J, and M over 5 time steps (t=0 to t=4). The table highlights values for A, J, and M in green, while B and E are consistently False (F).

t	B	E	A	J	M
0	F	F	F	F	F
1	F	T	F	T	F
2					
3					
4					

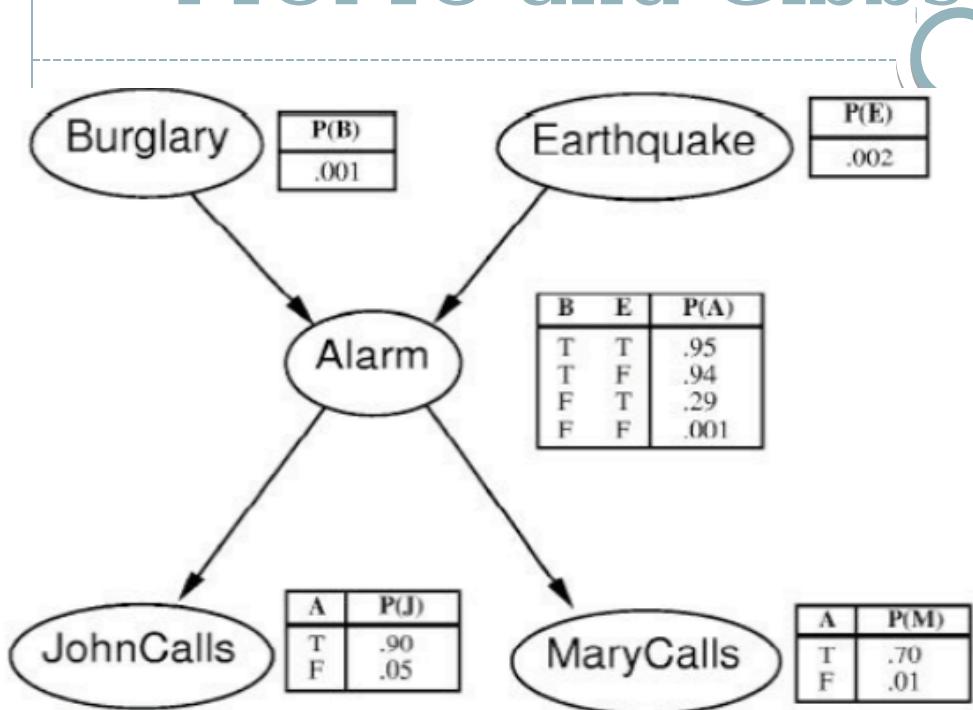
- $t=1$: campiono $P(M | A)$ direttamente dalla CPT:

- $A = F$

$$P(M = T | A = F) \propto 0.01$$

$$P(M = F | A = F) \propto 0.99$$

MCMC and Gibbs Sampler and PGM



t	B	E	A	J	M
0	F	F	F	F	F
1	F	T	F	T	F
2	F	T	T	T	T
3	T	F	T	F	T
4	T	F	T	F	F

- t=2: ripeto la procedura su B,E,A,J,M
- t=3: ripeto la procedura su B,E,A,J,M
- t=4: ripeto la procedura su B,E,A,J,M

Libri e Documentazioni di Riferimento

