Stefano Gariazzo

IFIC, Valencia (ES)
CSIC – Universitat de Valencia

gariazzo@ific.uv.es
http://ific.uv.es/~gariazzo/

Horizon 2020
European Union funding
for Research & Innovation

Autostati di massa dei neutrini e il loro ordine: un approccio Bayesiano

Basato su JCAP 03 (2018) 011

05/04/2018 - Incontri Fisica Alte Energie - Milano (IT)

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\,p(H|I)}{p(d|I)}$$

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\, p(H|I)}{p(d|I)}$$

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori: cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\, p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori: cosa sapevamo prima

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori:

cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\, p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Likelihood: $\mathcal{L}(\theta)$

fondata sui dati, assumendo H vera

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori:

cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\,p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Likelihood: $\mathcal{L}(\theta)$

fondata sui dati, assumendo H vera

Evidenza Bayesiana:

$$p(d|I) \equiv \sum_H p(d|H,I)\,p(H|I)$$

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori:

cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\,p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Likelihood: $\mathcal{L}(\theta)$

fondata sui dati, assumendo H vera

Evidenza Bayesiana:

$$p(d|I) \equiv \sum_H p(d|H,I)\,p(H|I)$$

Teorema di Bayes:

$$\text{p. posteriori} = \frac{\text{likelihood} \times \text{p. priori}}{\text{evidenza}}$$

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori:

cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\,p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Likelihood: $\mathcal{L}(\theta)$

fondata sui dati, assumendo H vera

Evidenza Bayesiana:

$$p(d|I) \equiv \sum_H p(d|H,I)\,p(H|I)$$

inferenza dei parametri

Teorema di Bayes:

$$\text{p. posteriori} = \frac{\text{likelihood} \times \text{p. priori}}{\text{evidenza}}$$

Il teorema di Bayes

su cosa si fonda la statistica Bayesiana?

data un'ipotesi H, i dati d, e qualche informazione I (assunta come vera):

$p(\theta)$

Probabilità a posteriori:

cosa sapremo dopo

Teorema di Bayes:

$$p(H|d,I) = \frac{p(d|H,I)\,p(H|I)}{p(d|I)}$$

$\pi(\theta)$

Probabilità a priori:

cosa sapevamo prima

Likelihood: $\mathcal{L}(\theta)$

fondata sui dati, assumendo H vera

Evidenza Bayesiana:

$$p(d|I) \equiv \sum_H p(d|H,I)\,p(H|I)$$

inferenza dei parametri

Teorema di Bayes:

$$\text{p. posteriori} = \frac{\text{likelihood} \times \text{p. priori}}{\text{evidenza}}$$

paragone di modelli

Bayesian evidence

"Evidenza Bayesiana"

$$\boxed{p(d|\mathcal{M}) = {\color{red}Z} = \sum_H p(d|H,I)\, p(H|I)}$$

somma su diverse ipotesi (discrete)
(sempre assumendo I vera)

Bayesian evidence

"Evidenza Bayesiana"

$$p(d|\mathcal{M}) = {\color{red}Z} = \int_{\Omega_\mathcal{M}} p(d|\theta,\mathcal{M})\, p(\theta|\mathcal{M})\, d\theta$$

integrale sulle ipotesi (continue) = valori dei parametri nel modello $\mathcal{M}$ (considerando $\mathcal{M}$ corretto)

Bayesian evidence

“Evidenza Bayesiana”

$$p(d|\mathcal{M}) = {\color{red}Z} = \int_{\Omega_\mathcal{M}} p(d|\theta,\mathcal{M})\,p(\theta|\mathcal{M})\,d\theta$$

integrale sulle ipotesi (continue) $=$ valori dei parametri nel modello $\mathcal{M}$ (considerando $\mathcal{M}$ corretto)

Cosa succede se abbiamo diversi modelli $\mathcal{M}_i$?

usare Z_i per comparare i modelli

Bayesian evidence

"Evidenza Bayesiana"

$$\boxed{p(d|\mathcal{M}) = Z = \int_{\Omega_{\mathcal{M}}} p(d|\theta,\mathcal{M})\, p(\theta|\mathcal{M})\, d\theta}$$

integrale sulle ipotesi (continue) = valori dei parametri nel modello $\mathcal{M}$ (considerando $\mathcal{M}$ corretto)

Cosa succede se abbiamo diversi modelli $\mathcal{M}_i$?

usare Z_i per comparare i modelli

P. a posteriori del modello:

$$\boxed{p(\mathcal{M}_i|d) \propto p(\mathcal{M}_i)\, Z_i} \longleftarrow$$

costante di proporzionalità dipende solo dai dati

$p(\mathcal{M}_i)$ p. a priori del modello

Fattore di Bayes

È migliore $\mathcal{M}_1$ o $\mathcal{M}_2$?

$$\boxed{\frac{p(\mathcal{M}_1|d)}{p(\mathcal{M}_2|d)} = B_{1,2}\,\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}}$$

Fattore di Bayes:

$$\boxed{B_{1,2} = \frac{Z_1}{Z_2} \quad \Rightarrow \quad \ln B_{1,2} = \ln Z_1 - \ln Z_2}$$

Fattore di Bayes

È migliore $\mathcal{M}_1$ o $\mathcal{M}_2$?

$$\boxed{\frac{p(\mathcal{M}_1|d)}{p(\mathcal{M}_2|d)} = B_{1,2}\,\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}}$$

Fattore di Bayes:

$$\boxed{B_{1,2} = \frac{Z_1}{Z_2} \quad \Rightarrow \quad \ln B_{1,2} = \ln Z_1 - \ln Z_2}$$

se p. a priori sono uguali $[p(\mathcal{M}_1) = p(\mathcal{M}_2)]$,
$B_{1,2}$ determina il modello preferito:

$B_{1,2} > 1$ $(\ln B_{1,2} > 0)$

$\boxed{\mathcal{M}_1 \text{ preferito}}$

$B_{1,2} < 1$ $(\ln B_{1,2} < 0)$

$\boxed{\mathcal{M}_2 \text{ preferito}}$

Scala di Jeffreys

[Trotta, arxiv:0803.4089]

Preferenza per il modello migliore:

$$\boxed{|B_{1,2}| : 1}$$

rilevanza del segnale secondo la scala di Jeffreys:

$\lvert \ln B_{1,2} \rvert$	Preferenza	probabilità	rilevanza
< 1.0	$\lesssim 3:1$	< 0.750	inconclusiva
$\in [1.0, 2.5]$	$(3-12):1$	< 0.923	debole
$\in [2.5, 5.0]$	$(12-150):1$	< 0.993	moderata
> 5.0	$> 150:1$	> 0.993	forte

Preferenza & rilevanza sempre valide

probabilità corretta solo in caso di due modelli con uguale p. a priori (per esempio, ordine delle masse dei neutrini: normale O inverso)

Rasoio di Occam

Cosa ci insegna il paragone di modelli in statistica Bayesiana?

Miglior modello è il bilanciamento migliore fra

Qualità del fit → Rasoio di Occam ← Predittività

la teoria più semplice che spiega i dati è preferita

modello con più parametri → fit migliore (di solito)

→ veramente servono tutti i parametri?

il fattore di Bayes penalizza i modelli più complessi del necessario!

Rasoio di Occam

Cosa ci insegna il paragone di modelli in statistica Bayesiana?

Miglior modello è il bilanciamento migliore fra

Qualità del fit

Predittività

Rasoio di Occam

la teoria più semplice che spiega i dati è preferita

e se usiamo gli stessi parametri ma cambimo la loro p. a priori?

Evidenza bayesiana dipende da p. a priori!

Fattore di Bayes penalizza p. a priori più estese del necessario!

Rasoio di Occam

Cosa ci insegna il paragone di modelli in statistica Bayesiana?

Miglior modello è il bilanciamento migliore fra

Qualità del fit

Predittività

Rasoio di Occam

la teoria più semplice che spiega i dati è preferita

e se usiamo gli stessi parametri ma cambimo la loro p. a priori?

Evidenza bayesiana dipende da p. a priori!

Fattore di Bayes penalizza p. a priori più estese del necessario!

Fattore di Bayes NON penalizza i modelli in cui ci sono parametri non vincolati dai dati

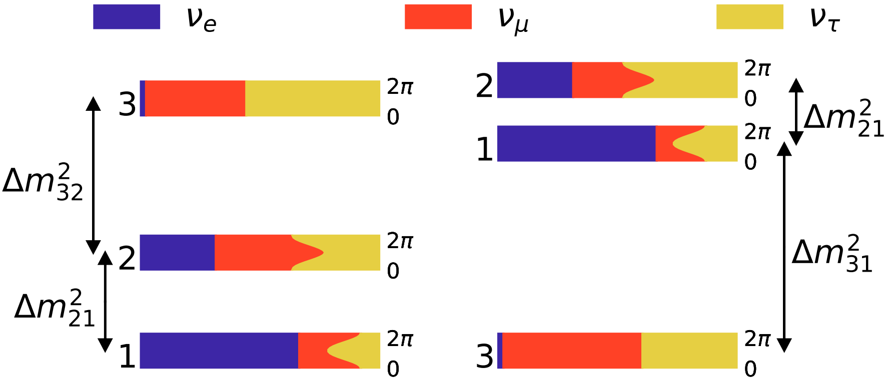
Masse dei neutrini

Ordine normale (NO)

$$m_1 < m_2 < m_3$$
$$\sum m_k \gtrsim 0.06 \text{ eV}$$

Ordine inverso (IO)

$$m_3 < m_1 < m_2$$
$$\sum m_k \gtrsim 0.1 \text{ eV}$$



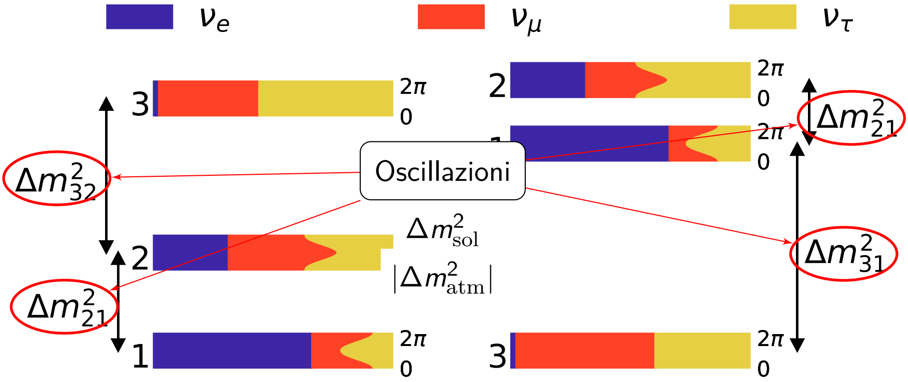
Masse dei neutrini

Ordine normale (NO)

$$m_1 < m_2 < m_3$$
$$\sum m_k \gtrsim 0.06 \text{ eV}$$

Ordine inverso (IO)

$$m_3 < m_1 < m_2$$
$$\sum m_k \gtrsim 0.1 \text{ eV}$$



Masse dei neutrini

Ordine normale (NO)

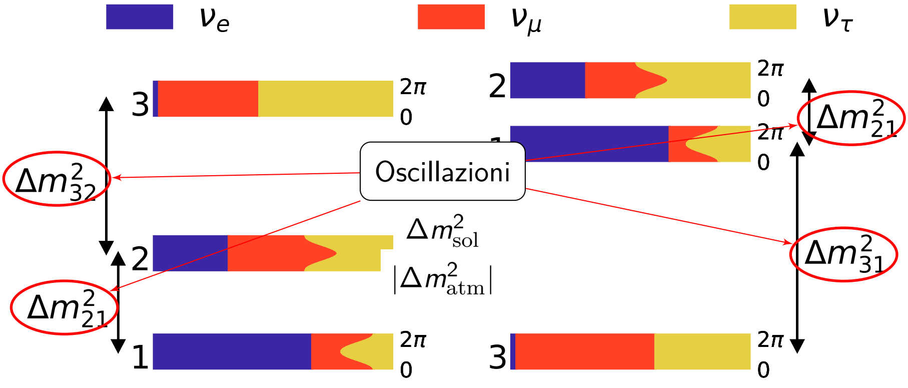
$m_1 < m_2 < m_3$

$\sum m_k \gtrsim 0.06$ eV

Ordine inverso (IO)

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV



Scala assoluta delle masse non nota!

Si può determinare l'ordine delle masse ponendo limiti su $\sum m_k$?

Limiti sulla scala assoluta delle masse dei neutrini

Effetti dei neutrini sullo spettro del decadimento β

Mainz/Troitsk: $m_{\nu_e} \lesssim 2$ eV

Katrin (atteso): $m_{\nu_e} \lesssim 0.2$ eV

$$m_{\nu_e}^2 = \sum_k |U_{ek}|^2 \, m_k^2$$

U_{ek} matrice PMNS

Limiti sulla scala assoluta delle masse dei neutrini

$\boxed{\text{Effetti dei neutrini sullo spettro del decadimento } \beta}$

Mainz/Troitsk: $m_{\nu_e} \lesssim 2$ eV

Katrin (atteso): $m_{\nu_e} \lesssim 0.2$ eV

$m^2_{\nu_e} = \sum_k |U_{ek}|^2 \, m^2_k$

U_{ek} matrice PMNS

(se il neutrino è particella di Majorana)

$\boxed{\text{Limiti dal doppio decadimento } \beta \text{ senza neutrini}}$

Misura $T^{0\nu}_{1/2}$, converti in $m_{\beta\beta}$ usando $m_e/m_{\beta\beta} = \mathcal{M}'^{\nu} \sqrt{G_{0\nu} \, T^{0\nu}_{1/2}}$

e infine $m_{\beta\beta} = \left|\sum_k e^{i\alpha_k} \, U^2_{ek} \, m_k\right|$

α_k fasi di Majorana

m_e massa elettrone,
$G_{0\nu}$ spazio fasi,
$\mathcal{M}'^{\nu}$ elemento di matrice

Limiti sulla scala assoluta delle masse dei neutrini

Effetti dei neutrini sullo spettro del decadimento β

Mainz/Troitsk: $m_{\nu_e} \lesssim 2$ eV

Katrin (atteso): $m_{\nu_e} \lesssim 0.2$ eV

$m^2_{\nu_e} = \sum_k |U_{ek}|^2 \, m^2_k$

U_{ek} matrice PMNS

(se il neutrino è particella di Majorana)

Limiti dal doppio decadimento β senza neutrini

Misura $T^{0\nu}_{1/2}$, converti in $m_{\beta\beta}$ usando $m_e/m_{\beta\beta} = \mathcal{M}'^{\nu} \sqrt{G_{0\nu} \, T^{0\nu}_{1/2}}$

e infine $m_{\beta\beta} = \left|\sum_k e^{i\alpha_k} \, U^2_{ek} \, m_k\right|$

α_k fasi di Majorana

m_e massa elettrone,
$G_{0\nu}$ spazio fasi,
$\mathcal{M}'^{\nu}$ elemento di matrice

Limiti cosmologici attraverso l'effetto delle masse dei neutrini

$\Rightarrow$ radiazione cosmica di fondo (CMB) ed espansione dell'universo

Al momento $\sum_k m_k \lesssim 0.1X$ eV, misure di singole m_k in futuro?

Cosa possiamo ricavare dai dati odierni?

1. [Hannestad, Schwetz, 2016]: preferenza estremamente debole (2:1, 3:2) per NO (cosmologia + fit di oscillazioni [Bergstrom et al., 2015]) approccio bayesiano;
2. [Gerbino et al, 2016]: preferenza estremamente debole (up to 3:2) per NO (solo cosmologia), approccio bayesiano;
3. [Simpson et al., 2017]: forte preferenza per NO (limiti cosmologici su $\sum m_\nu$ + misure di Δm^2_{21} e $|\Delta m^2_{31}|$) approccio bayesiano;
4. [Schwetz et al., 2017], "Comment on [Simpson et al., 2017]": effetto di p. a priori?
5. [Capozzi et al., 2017]: preferenza a 2σ per NO (cosmologia + fit di oscillazioni [Capozzi et al., 2016, aggiornamento 2017]) approccio frequentista;
6. [Caldwell et al., 2017] indicazione molto debole per NO (cosmologia + $\beta\beta0\nu$ + fit di oscillazioni [Esteban et al., 2016] riadattato) approccio bayesiano;
7. [Wang, Xia, 2017]: Fattore di Bayes NO vs IO non informativo (solo cosmologia).

Cosa possiamo ricavare dai dati odierni?

1. [Hannestad, Schwetz, 2016]: preferenza estremamente debole (2:1, 3:2) per NO (cosmologia + fit di oscillazioni [Bergstrom et al., 2015]) approccio bayesiano;
2. [Gerbino et al, 2016]: preferenza estremamente debole (up to 3:2) per NO (solo cosmologia), approccio bayesiano;
3. [Simpson et al., 2017]: forte preferenza per NO (limiti cosmologici su $\sum m_\nu$ + misure di Δm^2_{21} e $|\Delta m^2_{31}|$) approccio bayesiano;
4. [Schwetz et al., 2017], "Comment on [Simpson et al., 2017]": effetto di p. a priori?
5. [Capozzi et al., 2017]: preferenza a 2σ per NO (cosmologia + fit di oscillazioni [Capozzi et al., 2016, aggiornamento 2017]) approccio frequentista;
6. [Caldwell et al., 2017] indicazione molto debole per NO (cosmologia + $\beta\beta0\nu$ + fit di oscillazioni [Esteban et al., 2016] riadattato) approccio bayesiano;
7. [Wang, Xia, 2017]: Fattore di Bayes NO vs IO non informativo (solo cosmologia).

Parametrizzazioni, p. a priori e dati

[SG et al., JCAP 03 (2018) 011]

Oscillazioni di neutrini

$$\chi^2 = -2\log\mathcal{L}_{\rm osc}$$

dal fit globale

[de Salas et al, 2017]

Oscillazioni

Parametro	p. a priori
$\sin^2\theta_{12}$	0.1 – 0.6
$\sin^2\theta_{13}$	0.00 – 0.06
$\sin^2\theta_{23}$	0.25 – 0.75

Masse: vedi oltre!

[SG et al., JCAP 03 (2018) 011]

Dati $\beta\beta0\nu$

Likelihood approssimate come in [Caldwell et al, 2017], usando [Gerda, 2017] (Ge), [KamLAND-Zen, 2016], [EXO-200, 2014] (Xe)

Oscillazioni di neutrini

$\chi^2 = -2\log\mathcal{L}_{\rm osc}$ dal fit globale [de Salas et al, 2017]

$\beta\beta0\nu$		Oscillazioni	
Parametro	p. a priori	Parametro	p. a priori
α_2	$0 - 2\pi$	$\sin^2\theta_{12}$	0.1 – 0.6
α_3	$0 - 2\pi$	$\sin^2\theta_{13}$	0.00 – 0.06
$\mathcal{M}^{0\nu}_{^{76}Ge}$	4.07 – 4.87	$\sin^2\theta_{23}$	0.25 – 0.75
$\mathcal{M}^{0\nu}_{^{136}Xe}$	2.74 – 3.45		

Masse: vedi oltre!

Parametrizzazioni, p. a priori e dati

[SG et al., JCAP 03 (2018) 011]

Dati cosmologici	Dati $\beta\beta0\nu$	Oscillazioni di neutrini
Dati completi CMB (temperatura, polarizzazione) di [Planck, 2015], con modello ΛCDM come base	Likelihood approssimate come in [Caldwell et al, 2017], usando [Gerda, 2017] (Ge), [KamLAND-Zen, 2016], [EXO-200, 2014] (Xe)	$\chi^2 = -2\log\mathcal{L}_{\rm osc}$ dal fit globale [de Salas et al, 2017]

Cosmologia		$\beta\beta0\nu$		Oscillazioni	
Parametro	p. a priori	Parametro	p. a priori	Parametro	p. a priori
ω_b	0.019 – 0.025	α_2	$0 - 2\pi$	$\sin^2\theta_{12}$	0.1 – 0.6
ω_c	0.095 – 0.145	α_3	$0 - 2\pi$	$\sin^2\theta_{13}$	0.00 – 0.06
Θ_s	1.03 – 1.05	$\mathcal{M}^{0\nu}_{^{76}Ge}$	4.07 – 4.87	$\sin^2\theta_{23}$	0.25 – 0.75
τ	0.01 – 0.4	$\mathcal{M}^{0\nu}_{^{136}Xe}$	2.74 – 3.45		
n_s	0.885 – 1.04				
$\log(10^{10}A_s)$	2.5 – 3.7				

Masse: vedi oltre!

Parametrizzare le masse dei neutrini

[SG et al., JCAP 03 (2018) 011]

[Simpson et al, 2017]

usando m_1, m_2, m_3 (A)

[Caldwell et al, 2017]

usando m_{lightest}, Δm^2_{21}, $|\Delta m^2_{31}|$ (B)

intuizione dice: (B) è più vicino a osservabili! Meglio di (A)?

P. a priori deve essere lineare o logaritmica in m_k (m_{lightest})?

I dati possono dire chi è meglio, (A) o (B), lineare o log?

Parametrizzare le masse dei neutrini

[SG et al., JCAP 03 (2018) 011]

[Simpson et al, 2017]

usando m_1, m_2, m_3 (A)

[Caldwell et al, 2017]

usando m_{lightest}, Δm^2_{21}, $|\Delta m^2_{31}|$ (B)

intuizione dice: (B) è più vicino a osservabili! Meglio di (A)?

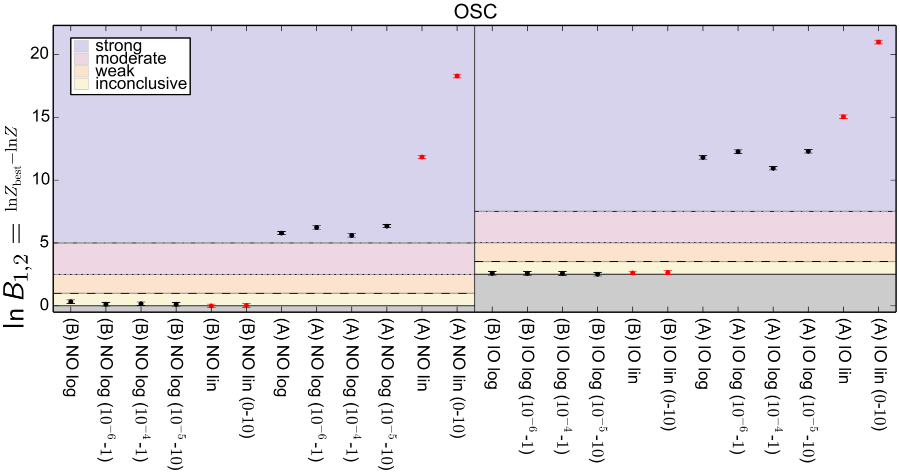
P. a priori deve essere lineare o logaritmica in m_k (m_{lightest})?

I dati possono dire chi è meglio, (A) o (B), lineare o log?

Caso A			Caso B		
Parameter	P. priori	Intervallo	Parametro	P. priori	Intervallo
m_1/eV	lineare log	$0 - 1$ $10^{-5} - 1$	$m_{\text{lightest}}/\text{eV}$	lineare log	$0 - 1$ $10^{-5} - 1$
m_2/eV	lineare log	$0 - 1$ $10^{-5} - 1$	$\Delta m^2_{21}/\text{eV}^2$	lineare	$5 \times 10^{-5} - 10^{-4}$
m_3/eV	lineare log	$0 - 1$ $10^{-5} - 1$	$\lvert\Delta m^2_{31}\rvert/\text{eV}^2$	lineare	$1.5 \times 10^{-3} - 3.5 \times 10^{-3}$

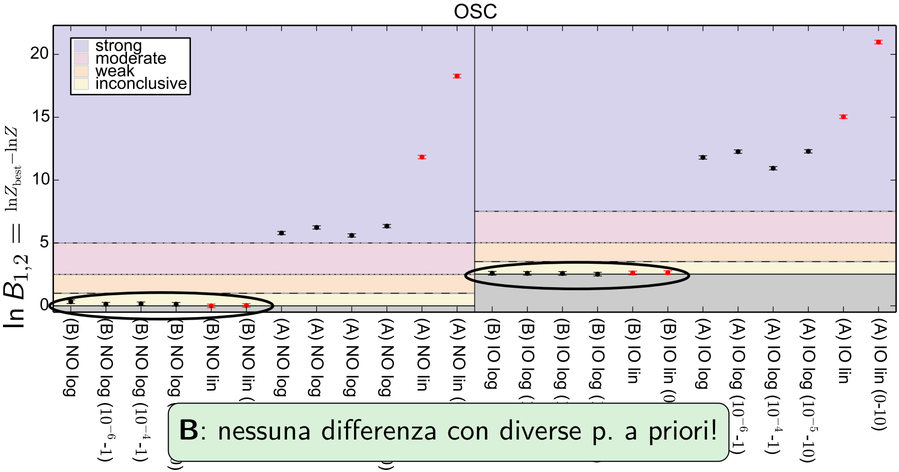
Paragoni fra parametrizzazioni e p. a priori

[SG et al., JCAP 03 (2018) 011]



Paragoni fra parametrizzazioni e p. a priori

[SG et al., JCAP 03 (2018) 011]

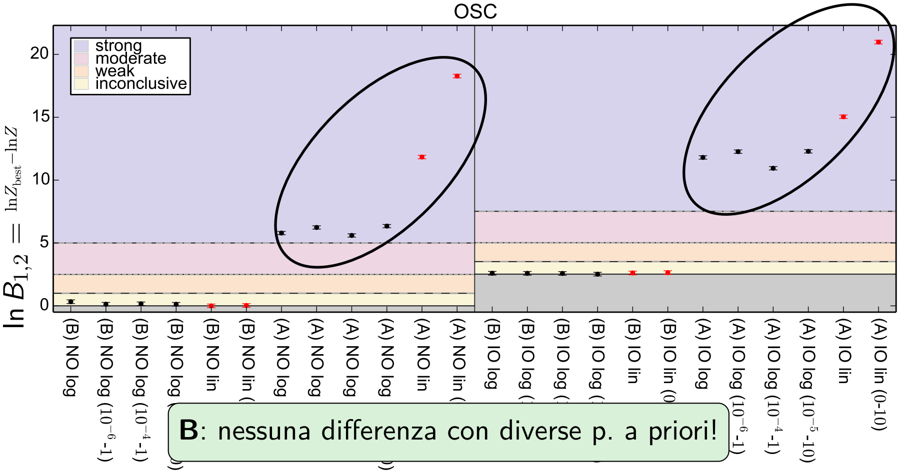


B: nessuna differenza con diverse p. a priori!

(nessuna penalità per parametro non vincolato, $m_{\rm lightest}$!)

Paragoni fra parametrizzazioni e p. a priori

[SG et al., JCAP 03 (2018) 011]



B: nessuna differenza con diverse p. a priori!

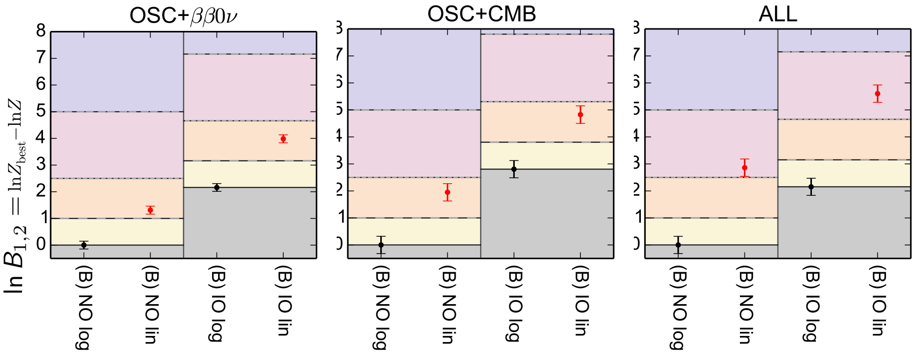
(nessuna penalità per parametro non vincolato, $m_{\rm lightest}$!)

A: sempre fortemente sfavorito!

(molto spazio parametri “sprecato”, no parametri non vincolati!)

Paragoni fra parametrizzazioni e p. a priori

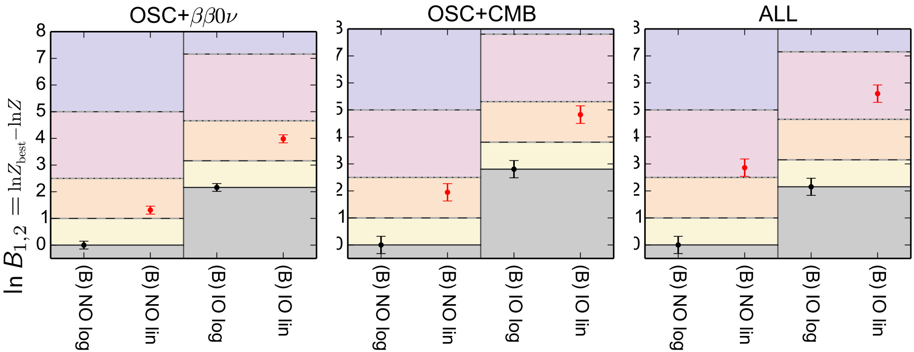
[SG et al., JCAP 03 (2018) 011]



compariamo lineare vs **logaritmico**

Paragoni fra parametrizzazioni e p. a priori

[SG et al., JCAP 03 (2018) 011]

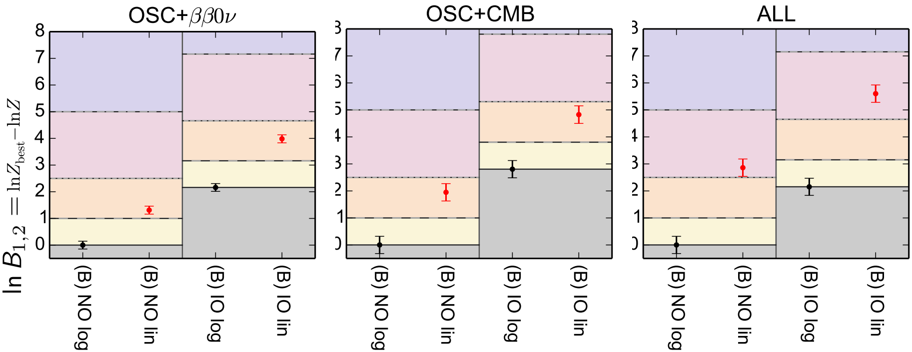


compariamo lineare vs **logaritmico**

log è
debolmente/moderatamente più efficiente

Paragoni fra parametrizzazioni e p. a priori

[SG et al., JCAP 03 (2018) 011]



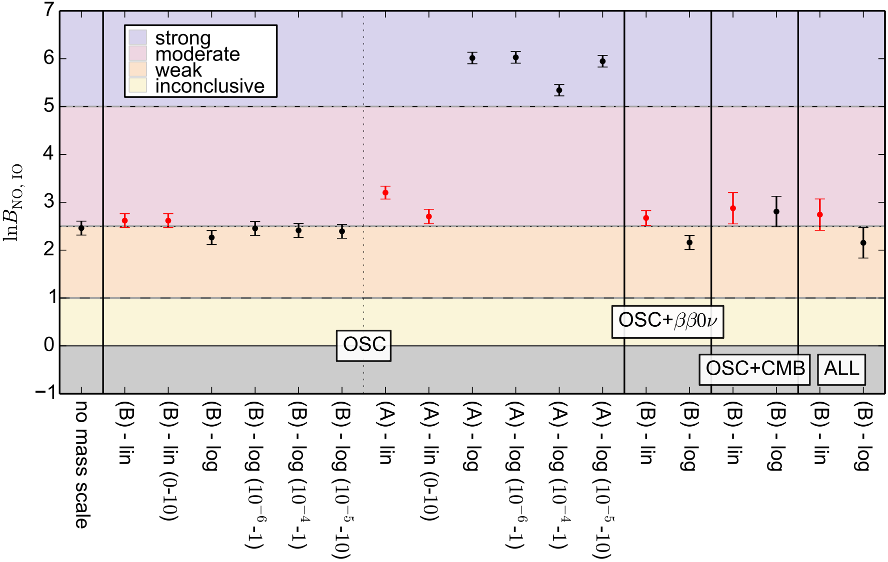
compariamo lineare vs **logaritmico**

log è
debolmente/moderatamente più efficiente

per riassumere: caso B, log è meglio!

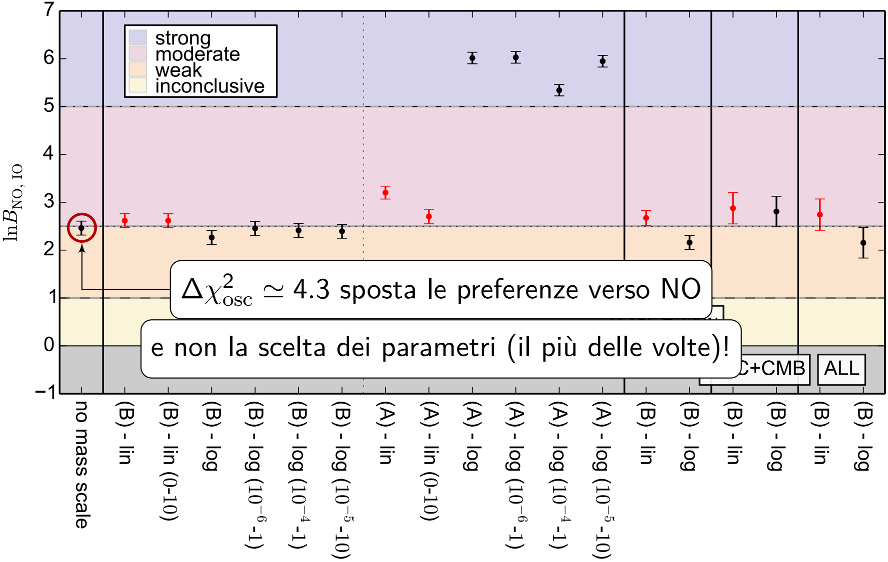
Paragoni fra NO e IO

[SG et al., JCAP 03 (2018) 011]



Paragoni fra NO e IO

[SG et al., JCAP 03 (2018) 011]

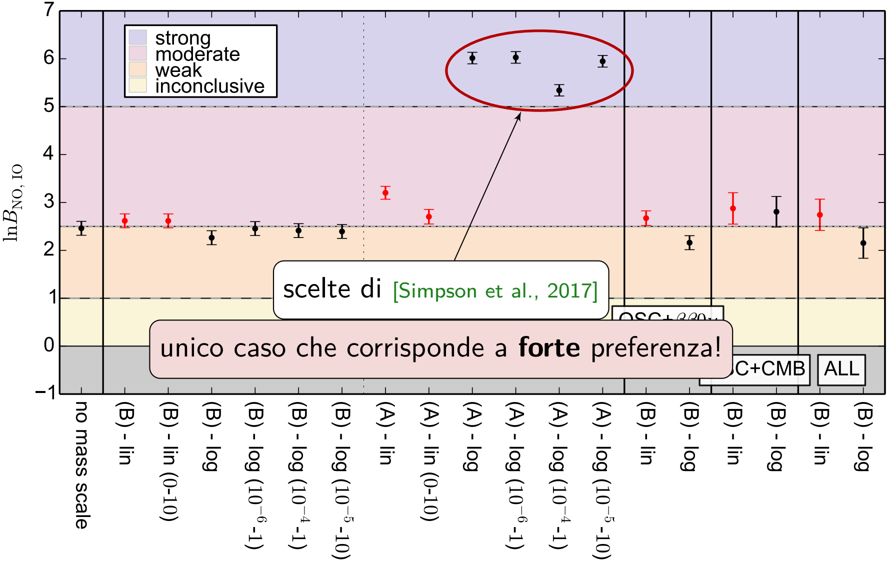


$\Delta\chi^2_{\mathrm{osc}} \simeq 4.3$ sposta le preferenze verso NO

e non la scelta dei parametri (il più delle volte)!

Paragoni fra NO e IO

[SG et al., JCAP 03 (2018) 011]



scelte di [Simpson et al., 2017]

unico caso che corrisponde a **forte** preferenza!

Conclusioni

1. Comparazione bayesiana di modelli
attraverso evidenza bayesiana/fattore di Bayes
per test robusti di parametrizzazioni
usando i dati

2. Attenzione alle scelte di p. a priori
(o di altre scelte soggettive)
che possono influenzare i risultati

3. i dati preferiscono moderatamente l'ordine
normale vs inverso per le masse dei neutrini

Conclusioni

1. Comparazione bayesiana di modelli
attraverso evidenza bayesiana/fattore di Bayes
per test robusti di parametrizzazioni
usando i dati

2. Attenzione alle scelte di p. a priori
(o di altre scelte soggettive)
che possono influenzare i risultati

3. i dati preferiscono moderatamente l'ordine
normale vs inverso per le masse dei neutrini

Grazie per l'attenzione!