Stefano Gariazzo

*IFIC, Valencia (ES)*
*CSIC – Universitat de Valencia*

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

# Neutrino mass ordering: current status

*What Bayesian model comparison*
*and available data can tell us*

VI Postgraduate Meeting On Theoretical Physics, Valencia - 29/11/2017

1. Basics of Bayesian statistics
   - Bayes' theorem
   - Bayesian model comparison
2. Constraining the neutrino mass ordering
   - Introducing the problem
   - Comparing models and mass orderings
3. Constraining the neutrino masses
4. Conclusions

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Bayes theorem:

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

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Bayes theorem:

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

Prior probability: what we knew before

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Posterior probability: what we know after

Bayes theorem:

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

Prior probability: what we knew before

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Posterior probability:

what we know after

Bayes theorem:

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

Prior probability:

what we knew before

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

sampling distribution of data, given that $H$ is true

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Posterior probability:

what we know after

Bayes theorem:

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

Prior probability:

what we knew before

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

sampling distribution of data, given that $H$ is true

Marginal likelihood:

or "Bayesian evidence",

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

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Posterior probability:

what we know after

Bayes theorem:

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

Prior probability:

what we knew before

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

sampling distribution of data, given that $H$ is true

Marginal likelihood:

or "Bayesian evidence",

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

Bayes theorem:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

# Bayes' theorem

Basic rule to deal with Bayesian probability!

given hypothesis $H$, data $d$, some information $I$ (true):

Bayes theorem:

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

Posterior probability:

what we know after → parameter inference

Prior probability:

what we knew before

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

sampling distribution of data, given that $H$ is true

Marginal likelihood:

or "Bayesian evidence",

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

Bayes theorem:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

# Bayes' theorem

**Basic rule to deal with Bayesian probability!**

given hypothesis $H$, data $d$, some information $I$ (true):

Bayes theorem:

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

Posterior probability: what we know after → parameter inference

Prior probability: what we knew before

Likelihood: $\mathcal{L}(\theta)$ sampling distribution of data, given that $H$ is true

Marginal likelihood: or "Bayesian evidence",

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

Bayes theorem:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

posterior → parameter inference

evidence → model comparison

# Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

sum over different (discrete) hypothesis
(given that $I$ is true)

# Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

integrate over all possible (continuous) parameters of model $\mathcal{M}$
(given that $\mathcal{M}$ is true)

# Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

integrate over all possible (continuous) parameters of model $\mathcal{M}$
(given that $\mathcal{M}$ is true)

What if there are several possible models $\mathcal{M}_i$?

use $Z_i$ to perform bayesian model comparison

# Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

integrate over all possible (continuous) parameters of model $\mathcal{M}$
(given that $\mathcal{M}$ is true)

What if there are several possible models $\mathcal{M}_i$?

use $Z_i$ to perform bayesian model comparison

warning: compare models given the same data!

# Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

integrate over all possible (continuous) parameters of model $\mathcal{M}$
(given that $\mathcal{M}$ is true)

What if there are several possible models $\mathcal{M}_i$?

use $Z_i$ to perform bayesian model comparison

warning: compare models given the same data!

Model posterior:

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

← proportional to constant that depends only on data

given model prior $p(\mathcal{M}_i)$

# Bayes factor

Posterior odds of $\mathcal{M}_1$ versus $\mathcal{M}_2$:

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

Bayes factor:

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

# Bayes factor

Posterior odds of $\mathcal{M}_1$ versus $\mathcal{M}_2$:

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

Bayes factor:

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

if priors are the same $[p(\mathcal{M}_1) = p(\mathcal{M}_2)]$,
$B_{1,2}$ tells which one is preferred:

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

$\mathcal{M}_1$ preferred

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

$\mathcal{M}_2$ preferred

# Bayes factor

Posterior odds of $\mathcal{M}_1$ versus $\mathcal{M}_2$:

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

Bayes factor:

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

if priors are the same $[p(\mathcal{M}_1) = p(\mathcal{M}_2)]$,
$B_{1,2}$ tells which one is preferred:

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

$\mathcal{M}_1$ preferred

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

$\mathcal{M}_2$ preferred

$|B_{1,2}|$ tells the odds in favor of preferred model

# Jeffreys' scale

[Trotta, arxiv:0803.4089]

odds in favor of the preferred model:

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

strength of preference according to Jeffreys' scale:

| $\|\ln B_{1,2}\|$ | Odds | probability | strength of evidence |
|---|---|---|---|
| $< 1.0$ | $\lesssim 3 : 1$ | $< 0.750$ | inconclusive |
| $\in [1.0, 2.5]$ | $(3 - 12) : 1$ | $< 0.923$ | weak |
| $\in [2.5, 5.0]$ | $(12 - 150) : 1$ | $< 0.993$ | moderate |
| $> 5.0$ | $> 150 : 1$ | $> 0.993$ | strong |

odds & strength always valid

probability correct given equal priors and that only two models are possible (see e.g. neutrino mass ordering: normal OR inverted)

# Occam's razor

what the Bayesian model comparison tells us?

Best model strikes optimum balance between

Quality of fit → Occam's razor ← Predictivity

the simplest theory that fits data is preferred

model with more parameters ⟶ better fit (usually)

→ are all the parameters needed?

Bayes factor penalizes unnecessarily complex models!

# Occam's razor

what the Bayesian model comparison tells us?

Best model strikes optimum balance between

Quality of fit → Occam's razor ← Predictivity

the simplest theory that fits data is preferred

what if we compare same model and different priors?

Bayesian evidence depends on priors!

Bayes factor penalizes unnecessarily wide priors!

# Occam's razor

what the Bayesian model comparison tells us?

Best model strikes optimum balance between

Quality of fit

Predictivity

Occam's razor

the simplest theory that fits data is preferred

what if we compare same model and different priors?

Bayesian evidence depends on priors!

Bayes factor penalizes unnecessarily wide priors!

Bayes factor DOES NOT penalize models with parameters that are unconstrained by the data

# Neutrino masses

**Normal ordering (NO)**

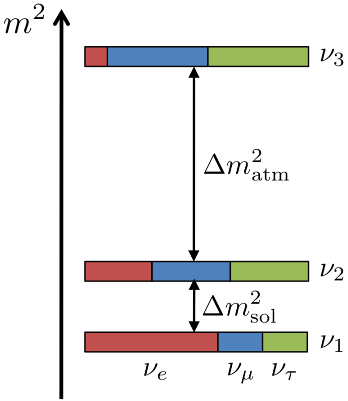
$m_1 < m_2 < m_3$

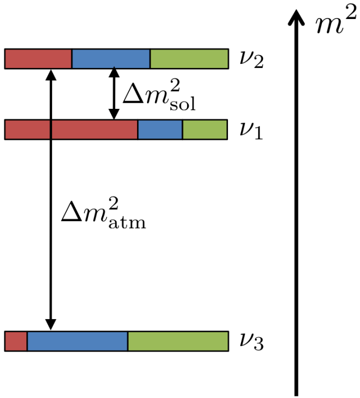
$\sum m_k \gtrsim 0.06$ eV

**Inverted ordering (IO)**

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV



# Neutrino masses

**Normal ordering (NO)**

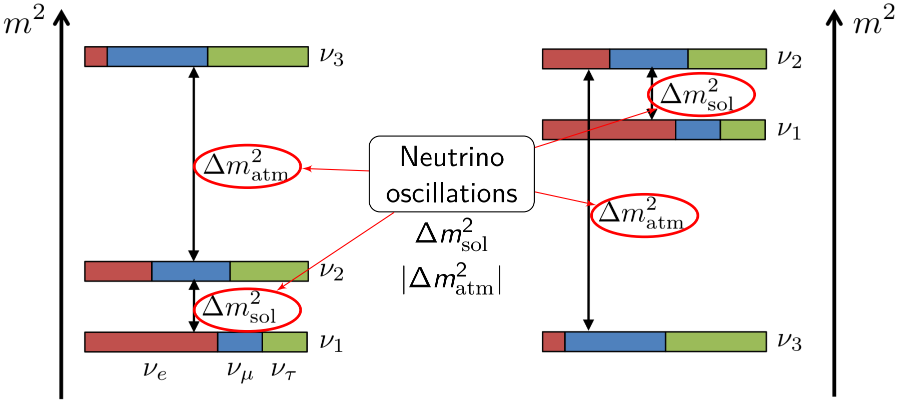
$m_1 < m_2 < m_3$

$\sum m_k \gtrsim 0.06$ eV

**Inverted ordering (IO)**

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV



# Neutrino masses

**Normal ordering (NO)**

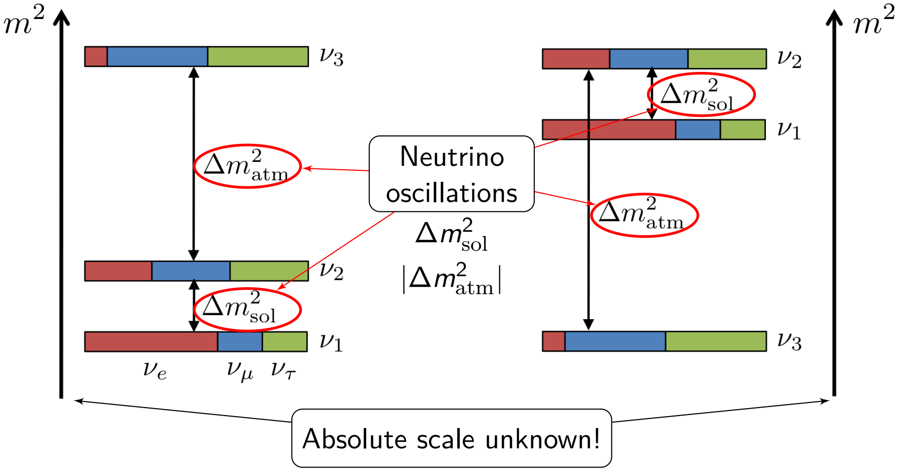
$m_1 < m_2 < m_3$

$\sum m_k \gtrsim 0.06$ eV

**Inverted ordering (IO)**

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV




Constrain mass ordering by constraining $\sum m_k$

# Constraining the absolute scale of neutrino masses

Neutrino effects on $\beta$ decay endpoint

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

Katrin, (expected) $m_{\nu_e} \lesssim 0.2$ eV

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

$U_{ek}$ mixing matrix

# Constraining the absolute scale of neutrino masses

Neutrino effects on $\beta$ decay endpoint

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

Katrin, (expected) $m_{\nu_e} \lesssim 0.2$ eV

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

$U_{ek}$ mixing matrix

(if neutrino is Majorana)

Constraints from neutrinoless double beta decay

Measure $T^{0\nu}_{1/2}$, convert into $m_{\beta\beta}$ using $m_e/m_{\beta\beta} = \mathcal{M}'^{\nu} \sqrt{G_{0\nu}\, T^{0\nu}_{1/2}}$

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

$\alpha_k$ Majorana phases

$m_e$ electron mass,
$G_{0\nu}$ phase space,
$\mathcal{M}'^{\nu}$ matrix element

# Constraining the absolute scale of neutrino masses

**Neutrino effects on $\beta$ decay endpoint**

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

Katrin, (expected) $m_{\nu_e} \lesssim 0.2$ eV

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

$U_{ek}$ mixing matrix

(if neutrino is Majorana)

**Constraints from neutrinoless double beta decay**

Measure $T^{0\nu}_{1/2}$, convert into $m_{\beta\beta}$ using $m_e/m_{\beta\beta} = \mathcal{M}'^{\nu}\sqrt{G_{0\nu}\,T^{0\nu}_{1/2}}$

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

$\alpha_k$ Majorana phases

$m_e$ electron mass,
$G_{0\nu}$ phase space,
$\mathcal{M}'^{\nu}$ matrix element

**Cosmological constraints from neutrino mass effects**

Mainly through CMB effects and free streaming

Currently can constrain $\sum_k m_k \lesssim 0.1X$ eV, single $m_k$ in future?

# Can current data tell us the neutrino mass ordering?

- [Hannestad, Schwetz, 2016]: extremely weak (2:1, 3:2) preference for NO (cosmology + [Bergstrom et al., 2015] neutrino oscillation fit) Bayesian approach;
- [Gerbino et al, 2016]: extremely weak (up to 3:2) preference for NO (cosmology only), Bayesian approach;
- [Simpson et al., 2017]: strong preference for NO (cosmological limits on $\sum m_\nu$ + constraints on $\Delta m^2_{21}$ and $|\Delta m^2_{31}|$) Bayesian approach;
- [Capozzi et al., 2017]: $2\sigma$ preference for NO (cosmology + [Capozzi et al., 2016, updated 2017] neutrino oscillation fit) frequentist approach;
- [Caldwell et al., 2017] very mild indication for NO (cosmology + neutrinoless double-beta decay + [Esteban et al., 2016] readapted oscillation results) Bayesian approach;
- [Wang, Xia, 2017]: Bayes factor NO vs IO is not informative (cosmology only).

## Can current data tell us the neutrino mass ordering?

- [Hannestad, Schwetz, 2016]: extremely weak (2:1, 3:2) preference for NO (cosmology + [Bergstrom et al., 2015] neutrino oscillation fit) Bayesian approach;
- [Gerbino et al, 2016]: extremely weak (up to 3:2) preference for NO (cosmology only), Bayesian approach;
- [Simpson et al., 2017]: strong preference for NO (cosmological limits on $\sum m_\nu$ + constraints on $\Delta m^2_{21}$ and $|\Delta m^2_{31}|$) Bayesian approach;
- [Capozzi et al., 2017]: $2\sigma$ preference for NO (cosmology + [Capozzi et al., 2016, updated 2017] neutrino oscillation fit) frequentist approach;
- [Caldwell et al., 2017] very mild indication for NO (cosmology + neutrinoless double-beta decay + [Esteban et al., 2016] readapted oscillation results) Bayesian approach;
- [Wang, Xia, 2017]: Bayes factor NO vs IO is not informative (cosmology only).

# Parameterizations, priors and data

[Gariazzo et al., in preparation]

Neutrino oscillations

full $\chi^2 = -2 \log \mathcal{L}_{\rm osc}$
from global fit

[de Salas et al, 2017]

Neutrino mixing

| Parameter | Prior |
|---|---|
| $\sin^2\theta_{12}$ | 0.1 – 0.6 |
| $\sin^2\theta_{13}$ | 0.00 – 0.06 |
| $\sin^2\theta_{23}$ | 0.25 – 0.75 |

Masses: see later!

# Parameterizations, priors and data

[Gariazzo et al., in preparation]

**$\beta\beta0\nu$ data**

Likelihood approximations as in [Caldwell et al, 2017], from [Gerda, 2017] (Ge), [KamLAND-Zen, 2016], [EXO-200, 2014] (Xe)

**Neutrino oscillations**

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

| $\beta\beta0\nu$ parameters | | Neutrino mixing | |
|---|---|---|---|
| Parameter | Prior | Parameter | Prior |
| $\alpha_2$ | $0 - 2\pi$ | $\sin^2\theta_{12}$ | 0.1 – 0.6 |
| $\alpha_3$ | $0 - 2\pi$ | $\sin^2\theta_{13}$ | 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 | | |

Masses: see later!

# Parameterizations, priors and data

[Gariazzo et al., in preparation]

**Cosmological data**

Full CMB temperature and polarization spectra from [Planck, 2015], working with ΛCDM model as basis

**$\beta\beta0\nu$ data**

Likelihood approximations as in [Caldwell et al, 2017], from [Gerda, 2017] (Ge), [KamLAND-Zen, 2016], [EXO-200, 2014] (Xe)

**Neutrino oscillations**

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

| Cosmological | | $\beta\beta0\nu$ parameters | | Neutrino mixing | |
|---|---|---|---|---|---|
| Parameter | Prior | Parameter | Prior | Parameter | Prior |
| $\omega_b$ | 0.019 – 0.025 | $\alpha_2$ | $0 - 2\pi$ | $\sin^2\theta_{12}$ | 0.1 – 0.6 |
| $\omega_c$ | 0.095 – 0.145 | $\alpha_3$ | $0 - 2\pi$ | $\sin^2\theta_{13}$ | 0.00 – 0.06 |
| $\Theta_s$ | 1.03 – 1.05 | $\mathcal{M}^{0\nu}_{76Ge}$ | 4.07 – 4.87 | $\sin^2\theta_{23}$ | 0.25 – 0.75 |
| $\tau$ | 0.01 – 0.4 | $\mathcal{M}^{0\nu}_{136Xe}$ | 2.74 – 3.45 | | |
| $n_s$ | 0.885 – 1.04 | | | | |
| $\log(10^{10}A_s)$ | 2.5 – 3.7 | | | | |

Masses: see later!

# Modeling neutrino masses

[Gariazzo et al., in preparation]

[Simpson et al, 2017]

using $m_1$, $m_2$, $m_3$ (A)

[Caldwell et al, 2017]

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

intuition says: (B) is closer to observable quantities! Better than (A)?

Should we use linear or logarithmic priors on $m_k$ ($m_{\rm lightest}$)?

Can data help to select (A) or (B), linear or log?

# Modeling neutrino masses

[Gariazzo et al., in preparation]

[Simpson et al, 2017]

using $m_1$, $m_2$, $m_3$ (A)

[Caldwell et al, 2017]

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

intuition says: (B) is closer to observable quantities! Better than (A)?

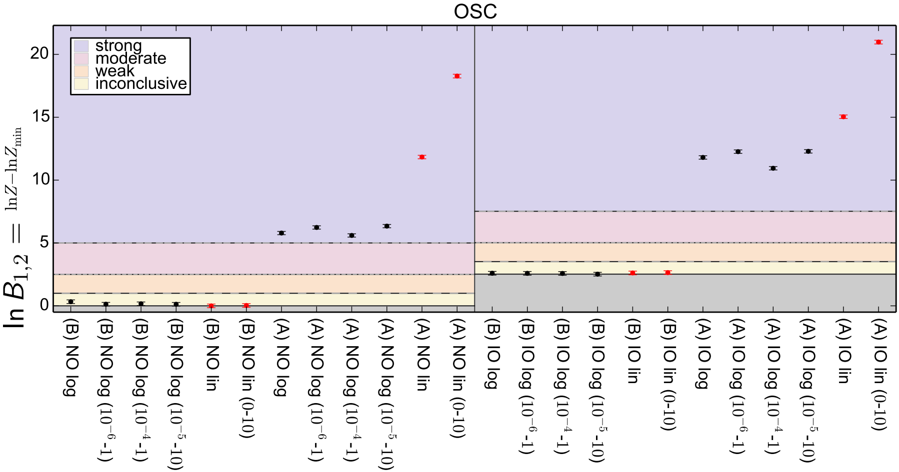
Should we use linear or logarithmic priors on $m_k$ ($m_{\rm lightest}$)?

Can data help to select (A) or (B), linear or log?

| Model A | | | Model B | | |
|---|---|---|---|---|---|
| Parameter | Prior | Range | Parameter | Prior | Range |
| $m_1/eV$ | linear<br>log | $0-1$<br>$10^{-5}-1$ | $m_{\rm lightest}/{\rm eV}$ | linear<br>log | $0-1$<br>$10^{-5}-1$ |
| $m_2/eV$ | linear<br>log | $0-1$<br>$10^{-5}-1$ | $\Delta m^2_{21}/{\rm eV}^2$ | linear | $5\times10^{-5}-10^{-4}$ |
| $m_3/eV$ | linear<br>log | $0-1$<br>$10^{-5}-1$ | $|\Delta m^2_{31}|/{\rm eV}^2$ | linear | $1.5\times10^{-3}-3.5\times10^{-3}$ |

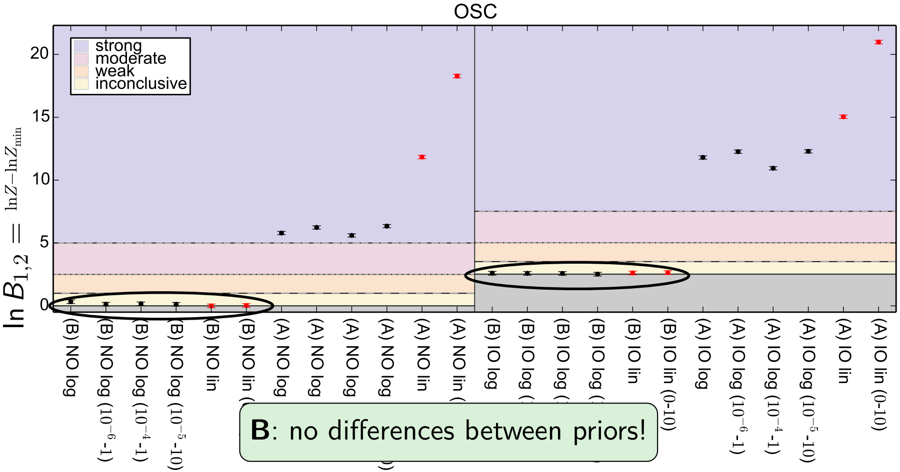
# Comparing models

[Gariazzo et al., in preparation]



# Comparing models

[Gariazzo et al., in preparation]

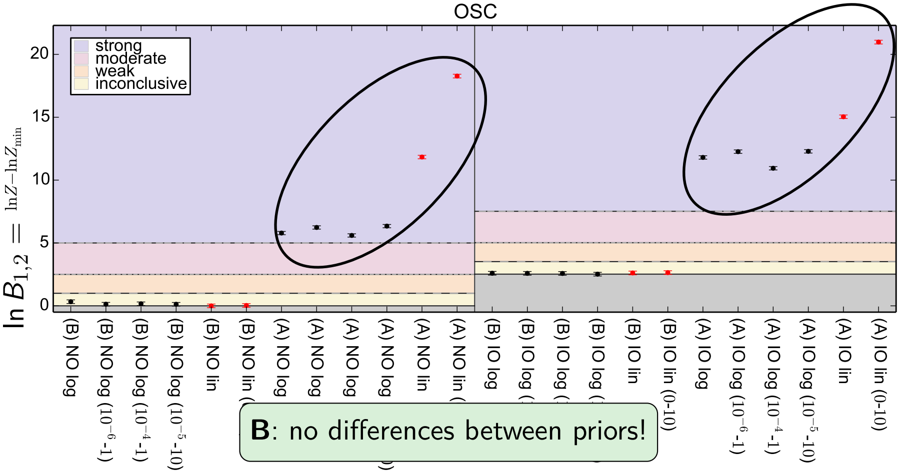



**B**: no differences between priors!

(no penalty for unconstrained parameter $m_{\text{lightest}}$!)

# Comparing models

[Gariazzo et al., in preparation]




**B**: no differences between priors!

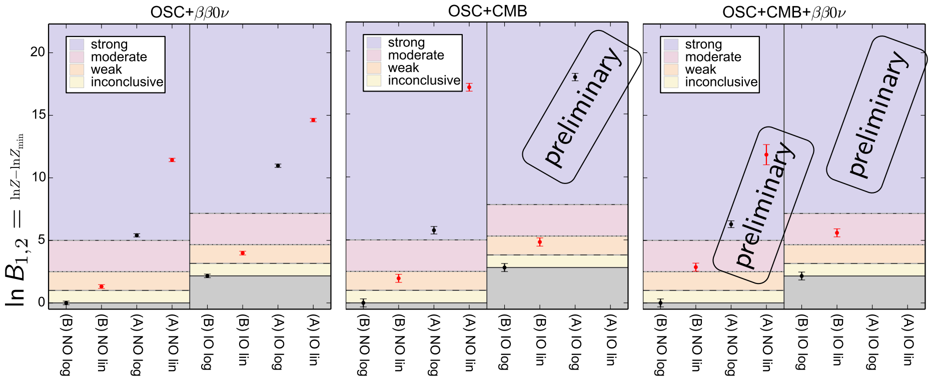
(no penalty for unconstrained parameter $m_{\text{lightest}}$!)

**A**: always strongly disfavored!

(waste of parameter space, no unconstrained parameter due to $\Delta m^2_{i1}$!)

# Comparing models

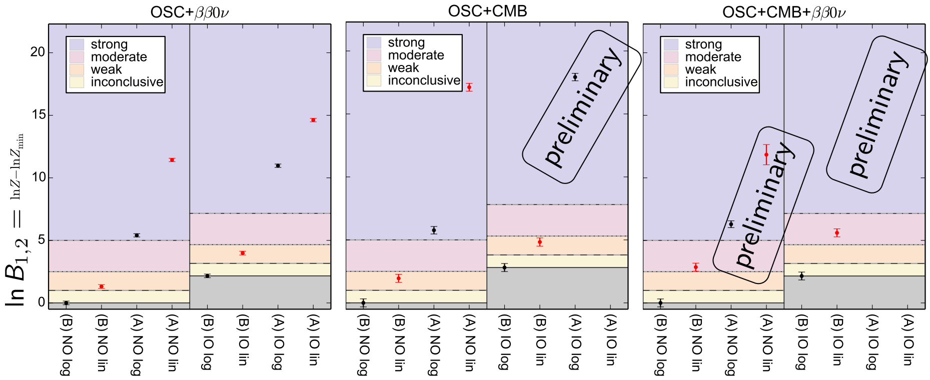
[Gariazzo et al., in preparation]




compare linear versus **logarithmic**

# Comparing models

[Gariazzo et al., in preparation]

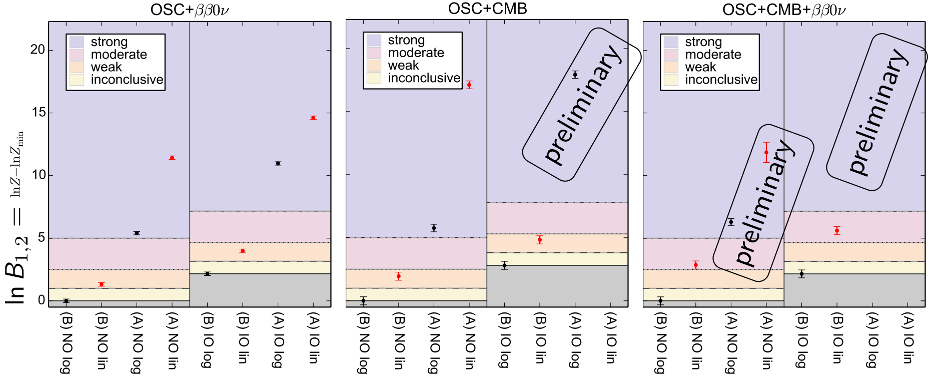



compare linear versus **logarithmic**

$\ln B_{\log,\,\mathrm{lin}}$ prefers **log** priors:
weakly-to-moderately for model B, moderately-to-strongly for model A

# Comparing models

[Gariazzo et al., in preparation]



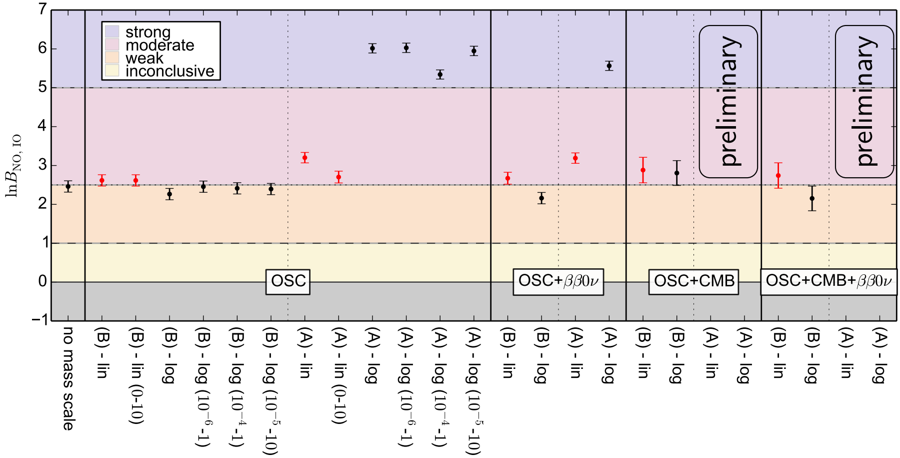

compare linear versus **logarithmic**

$\ln B_{\log,\,\mathrm{lin}}$ prefers **log** priors:
weakly-to-moderately for model B, moderately-to-strongly for model A

summary: model B, log priors is better!

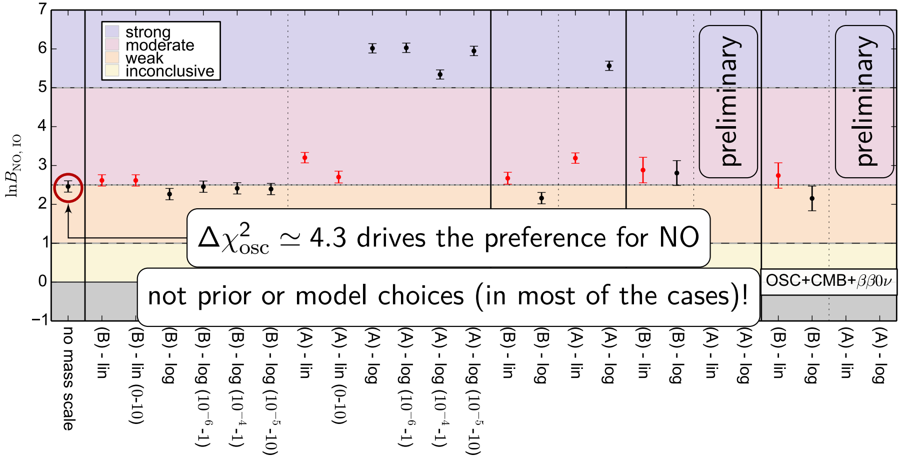
# Comparing the mass orderings

[Gariazzo et al., in preparation]



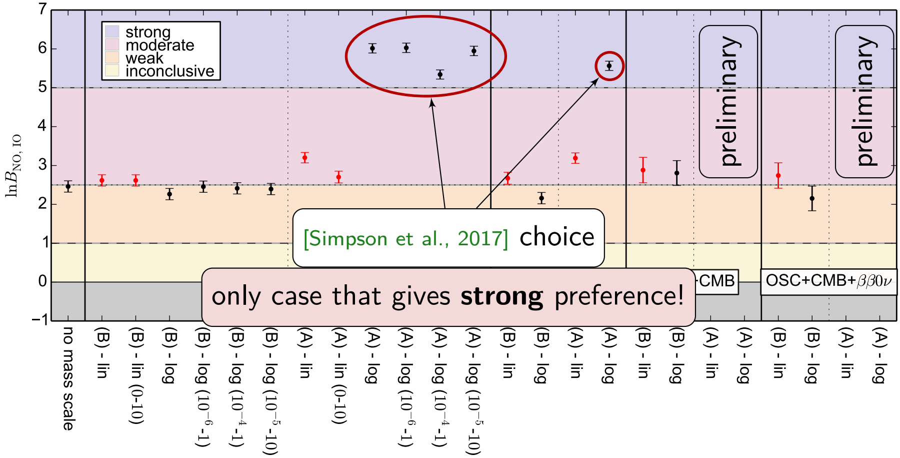
# Comparing the mass orderings

[Gariazzo et al., in preparation]



# Comparing the mass orderings

[Gariazzo et al., in preparation]

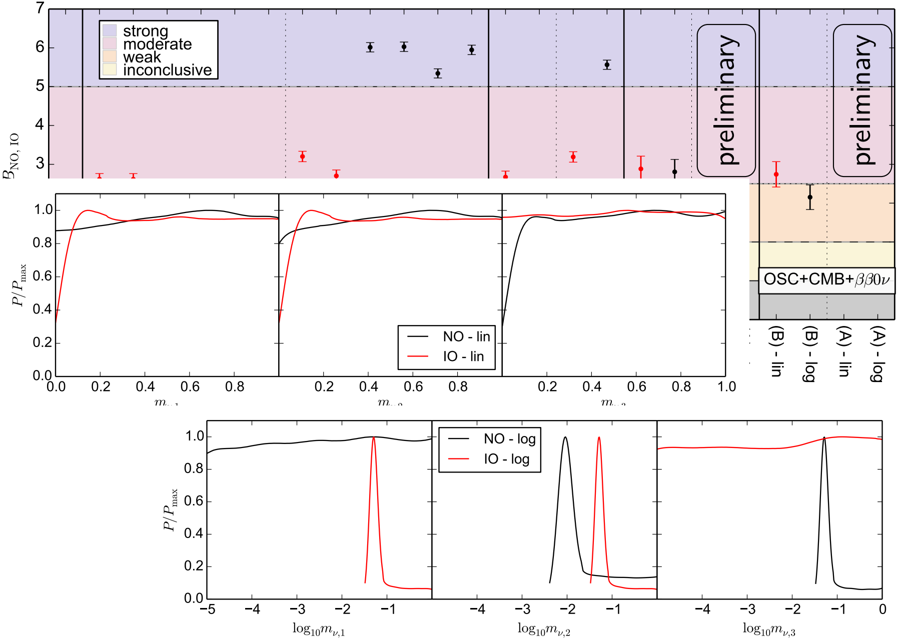



[Simpson et al., 2017] choice

only case that gives **strong** preference!

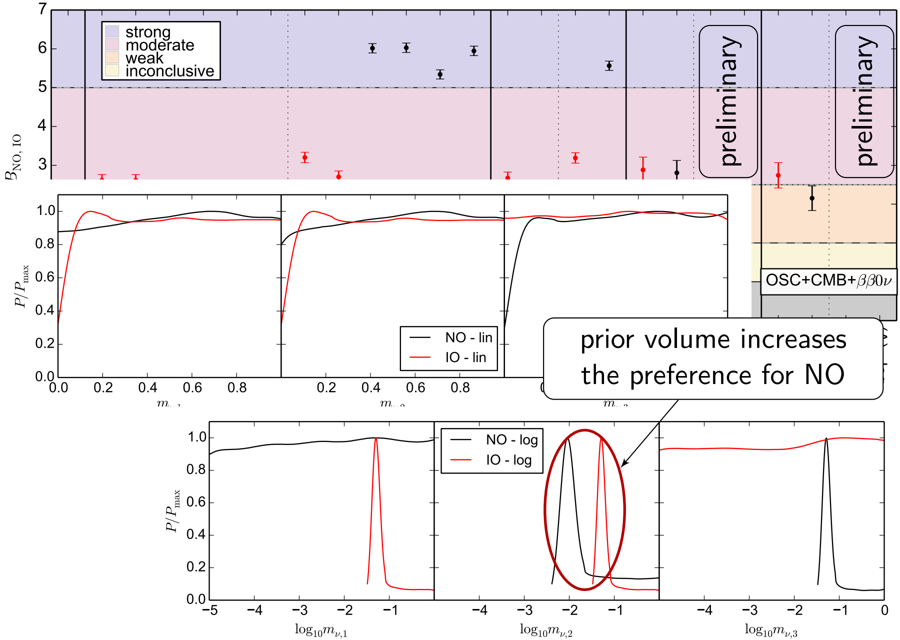
# Comparing the mass orderings

[Gariazzo et al., in preparation]



# Comparing the mass orderings

[Gariazzo et al., in preparation]



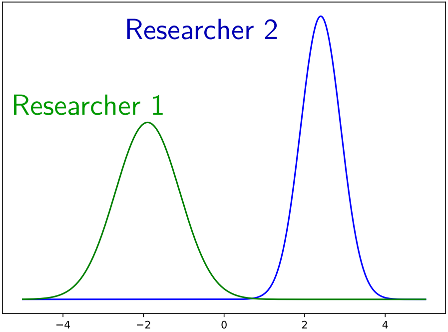

prior volume increases the preference for NO

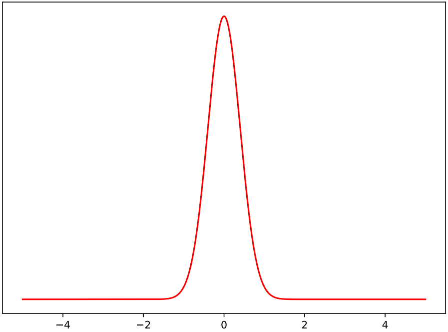
# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Prior

Likelihood



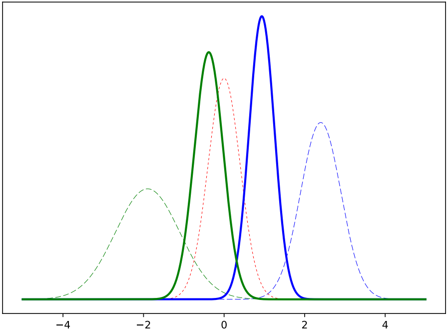

What each researcher knew before the experiment

The result of the experiment

# Bayes theorem in action

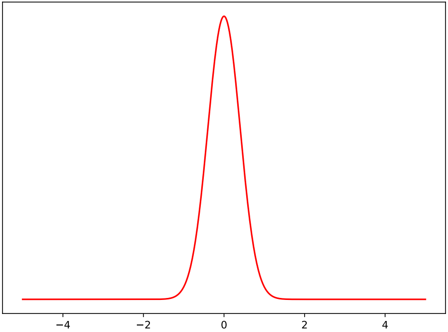
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior




What each researcher knows after the experiment

Likelihood




The result of the experiment

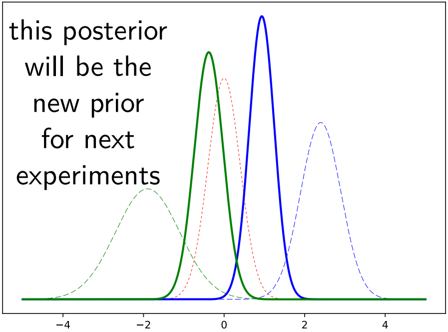
Posterior depends on prior!

# Bayes theorem in action

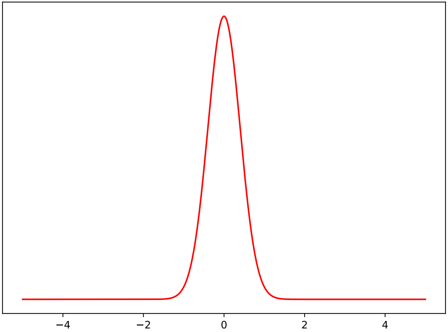
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior

Likelihood




What each researcher knows after the experiment




The result of the experiment

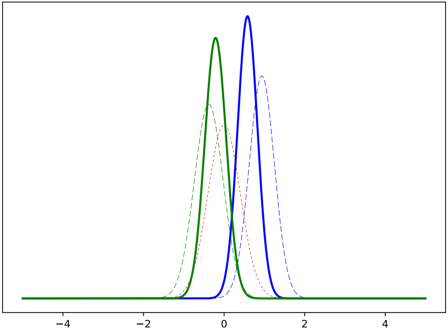
Posterior depends on prior!

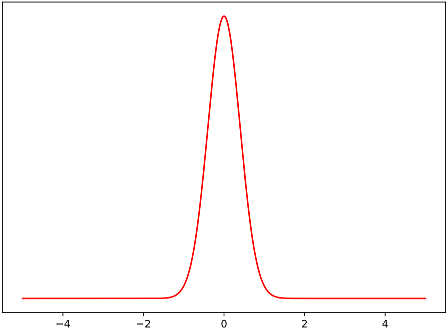
# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior

Likelihood




What each researcher knows after the second experiment

The result of the experiment

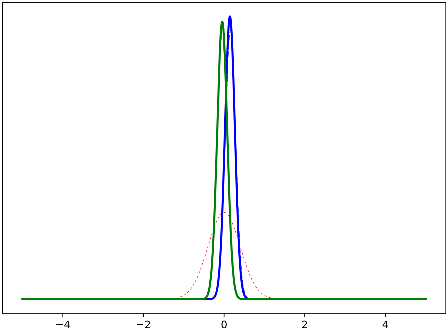
Remember:
$\sigma_N^2 = \sigma^2/N$

Posterior depends on prior!

# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior

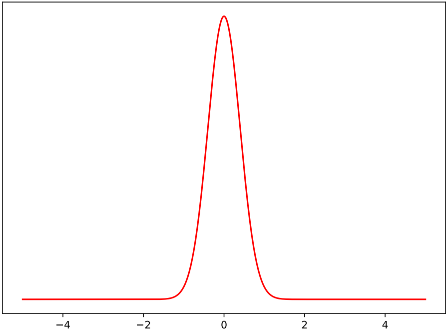



What each researcher knows after 10 experiments

Remember:

$\sigma_N^2 = \sigma^2/N$

Likelihood

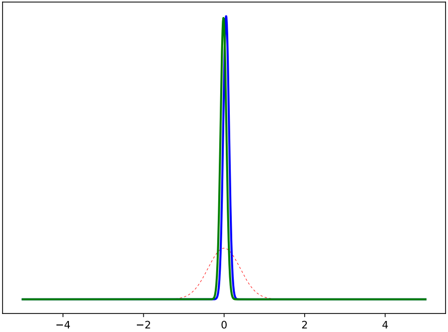



The result of the experiment

# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior

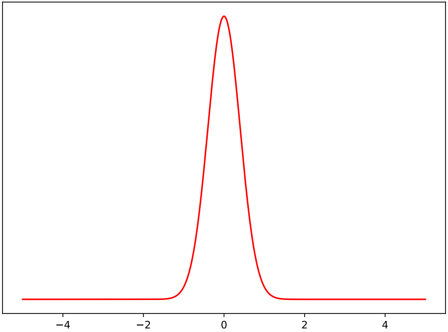



What each researcher knows after 30 experiments

Remember:

$\sigma_N^2 = \sigma^2/N$

Likelihood

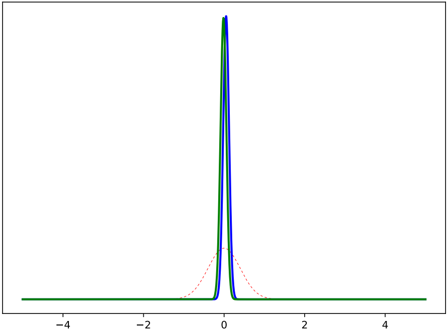



The result of the experiment

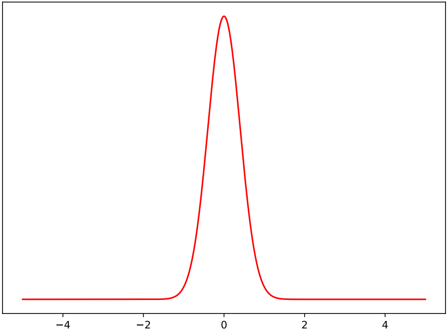
# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior




Likelihood

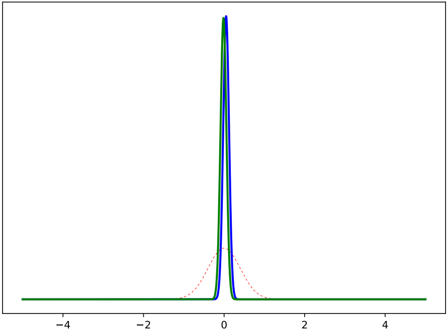



Knowledge converges using information from experiments

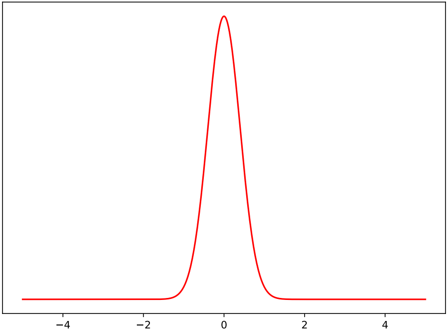
# Bayes theorem in action

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior




Likelihood

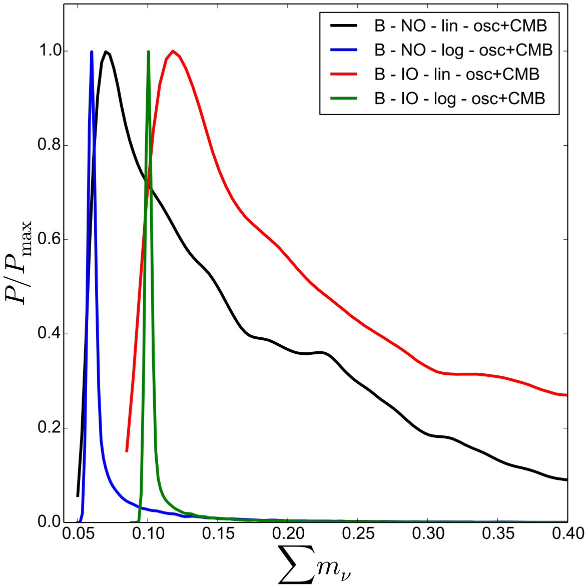



Knowledge converges using information from experiments

Prior dependence (subjectivity) only if not enough information in data!

# The role of priors: $\sum m_\nu$

[Gariazzo et al., in preparation]




showing model B only (only 1 parameter changes)

same for model A, but amplified (3 parameters change!)

logarithmic prior corresponds to $1/m_k$ probability!

↓

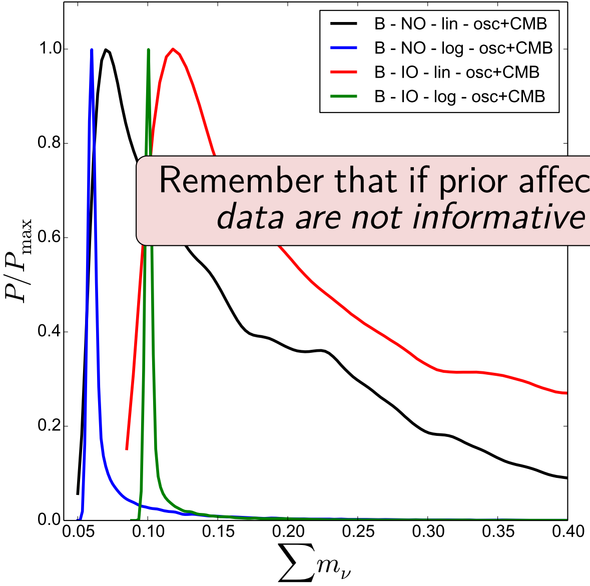
more importance to smaller masses

↓

limits closer to minimum allowed value of $\sum m_\nu$

# The role of priors: $\sum m_\nu$

[Gariazzo et al., in preparation]




Remember that if prior affects posterior,
*data are not informative enough!*

showing model B only (only 1 parameter changes)

same for model A, ...ied (3 ...change!)

logarithmic prior corresponds to $1/m_k$ probability!

↓

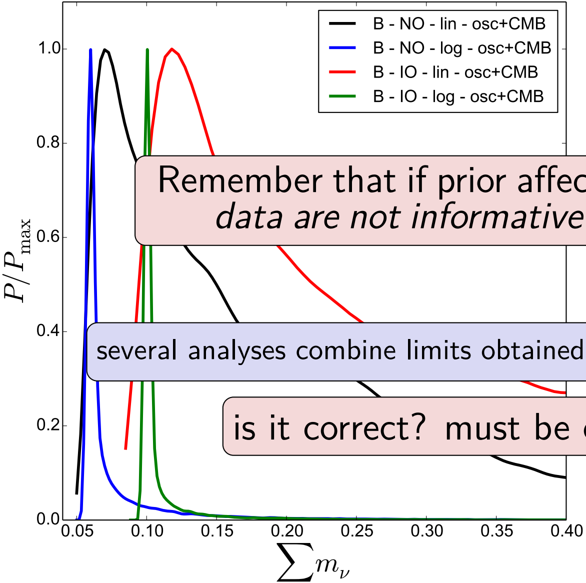
more importance to smaller masses

↓

limits closer to minimum allowed value of $\sum m_\nu$

# The role of priors: $\sum m_\nu$

[Gariazzo et al., in preparation]




showing model B only (only 1 parameter changes)

same for model A, ...ied (3 ...hange!)

logarithmic prior corresponds to ...ty!

...re importance ...maller masses

↓

limits closer to minimum allowed value of $\sum m_\nu$

Remember that if prior affects posterior, *data are not informative enough!*

several analyses combine limits obtained with different priors

is it correct? must be careful!

# Conclusions

1 Bayesian model comparison
through Bayesian evidence/Bayes factor
to robustly test models/priors against data

2 Be careful with the effects of prior
(or of other subjective choices)
on the results of your calculations

3 data only weakly/moderately prefer normal
versus inverted neutrino mass ordering

# Conclusions

1

Bayesian model comparison
through Bayesian evidence/Bayes factor
to robustly test models/priors against data

2

Be careful with the effects of prior
(or of other subjective choices)
on the results of your calculations

3

data only weakly/moderately prefer normal
versus inverted neutrino mass ordering

**Thank you for the attention!**