Stefano Gariazzo

IFIC, Valencia (ES)
CSIC – Universitat de Valencia

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

Bayesian model comparison applied to neutrino masses and their ordering

Based on arxiv:1801.04946

26/01/2018 - Seminar at IFIC - Valencia (ES)

Probability

[Trotta, arxiv:0803.4089]

What is probability?

a frequency

"the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions"

another subtle point: "randomness" of the trial series

what is really "random"?

do we properly know the initial state (and do not cheat)?

Probability

[Trotta, arxiv:0803.4089]

What is probability?

a frequency

"the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions"

another subtle point: "randomness" of the trial series

what is really "random"?

do we properly know the initial state (and do not cheat)?

Probability

[Trotta, arxiv:0803.4089]

What is probability?

a frequency

"the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions"

another subtle point:
"randomness" of the trial series

what is really "random"?

do we properly know the initial state (and do not cheat)?

Probability

What is probability?

a frequency

"the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions"

another subtle point: "randomness" of the trial series

what is really "random"?

do we properly know the initial state (and do not cheat)?

a degree of belief

"probability is a measure of the degree of belief about a preposition"

Probability

[Trotta, arxiv:0803.4089]

What is probability?

a frequency

"the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions"

another subtle point: "randomness" of the trial series

what is really "random"?

do we properly know the initial state (and do not cheat)?

a degree of belief

"probability is a measure of the degree of belief about a preposition"

Advantages:

- recovers frequentist on the long run;
- can be applied when frequentist cannot;
- no need to assume a distribution of possible data;
- deals effortlessly with nuisance parameters (*marginalization*);
- relies on *prior information*.

Bayes' theorem

how 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

how to deal with Bayesian probability?

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

Bayes theorem:

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

$\pi(\theta)$

Prior probability:

what we knew before

Bayes' theorem

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

Bayes theorem:

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

$\pi(\theta)$

Prior probability:

what we knew before

Bayes' theorem

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

Bayes theorem:

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

$\pi(\theta)$

Prior probability:

what we knew before

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

sampling distribution of data, given that H is true

Bayes' theorem

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

Bayes theorem:

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

$\pi(\theta)$

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

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

Bayes theorem:

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

$\pi(\theta)$

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

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

↓

parameter inference

Bayes theorem:

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

$\pi(\theta)$

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

how to deal with Bayesian probability?

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

$p(\theta)$

Posterior probability:

what we know after

Bayes theorem:

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

$\pi(\theta)$

Prior probability:

what we knew before

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

sampling distribution of data, given that H is true

Marginal likelihood:

or "Bayesian evidence",

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

parameter inference

Bayes theorem:

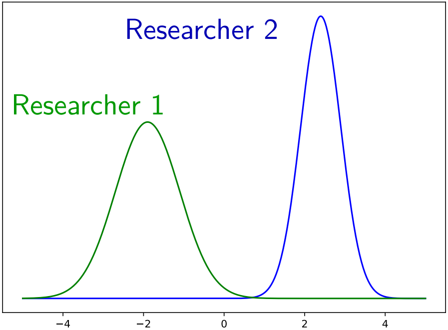
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

model comparison

Bayes theorem in action

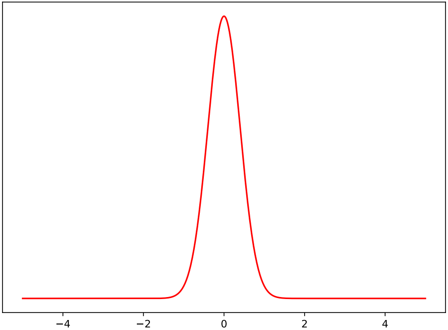
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Prior



What each researcher knew before the experiment

Likelihood

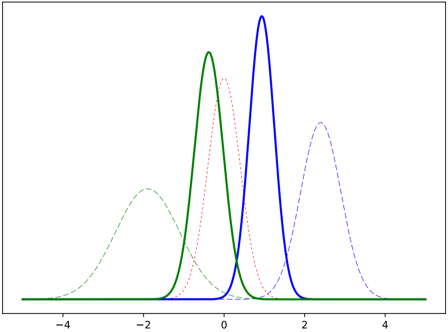


The result of the experiment

Bayes theorem in action

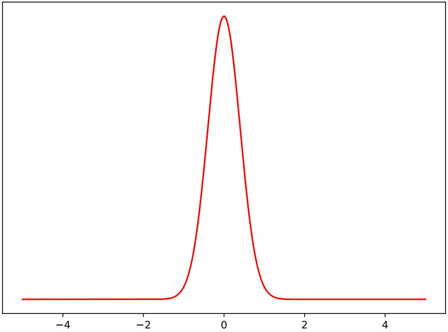
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior



What each researcher knows after the experiment

Likelihood



The result of the experiment

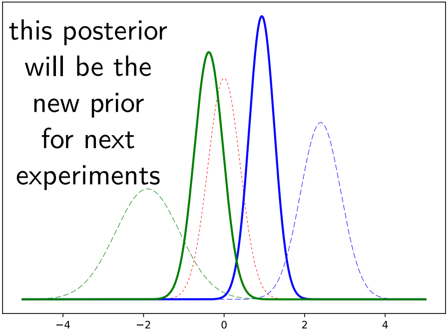
Posterior depends on prior!

Bayes theorem in action

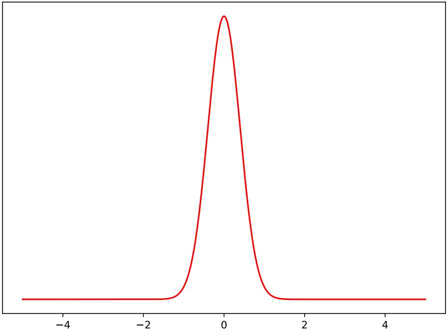
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior

Likelihood



What each researcher
knows after the experiment



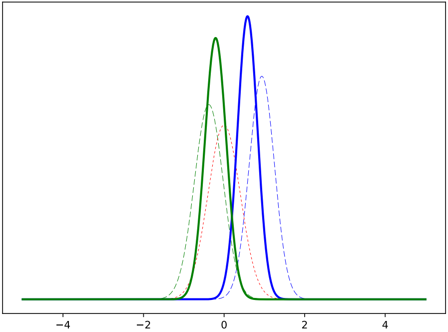
The result of the experiment

Posterior depends on prior!

Bayes theorem in action

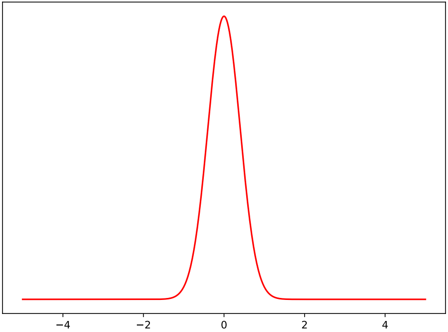
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior



What each researcher knows after the second experiment

Likelihood



The result of the experiment

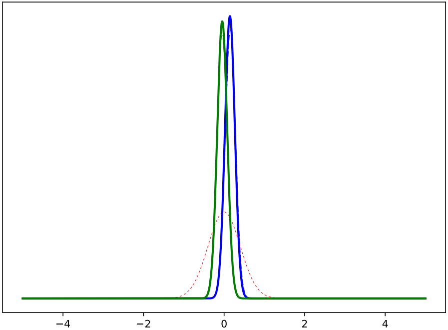
Remember:
$\sigma_N^2 = \sigma^2/N$

Posterior depends on prior!

Bayes theorem in action

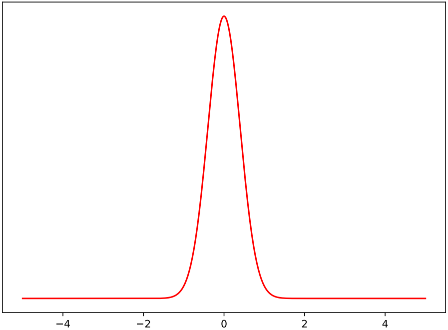
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior



What each researcher knows after 10 experiments

Likelihood



The result of the experiment

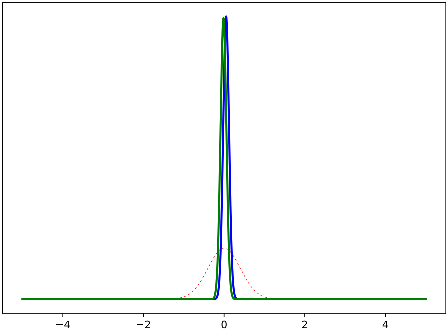
Remember:

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

Bayes theorem in action

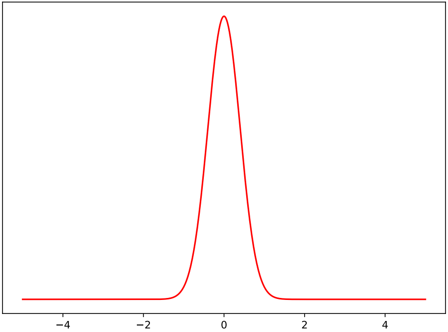
$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Posterior



What each researcher knows after 30 experiments

Likelihood



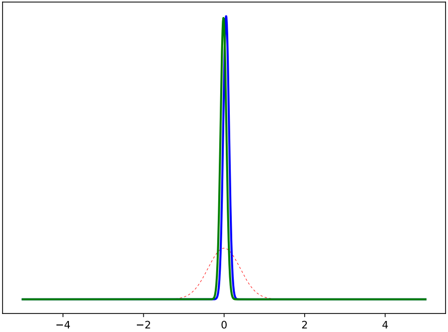
The result of the experiment

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

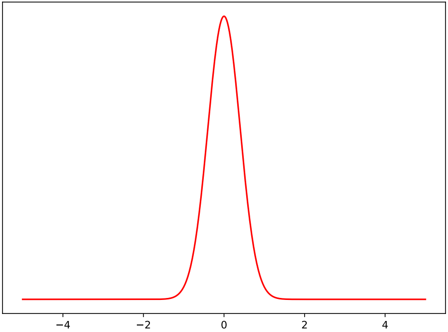
Bayes theorem in action

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

Posterior



Likelihood

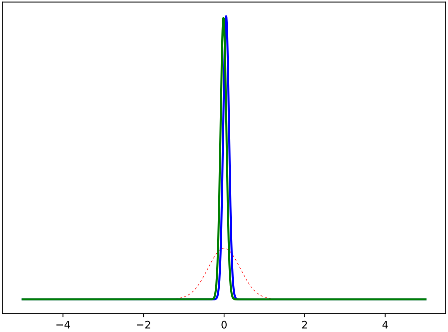


Knowledge converges using information from experiments

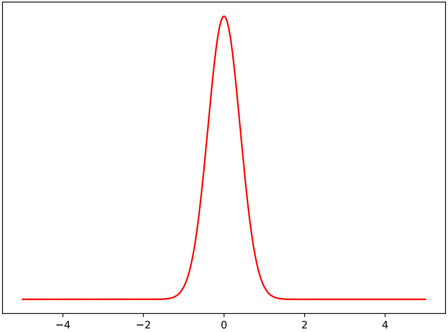
Bayes theorem in action

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

Posterior



Likelihood



Knowledge converges using information from experiments

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

Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

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

Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

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

Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

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"

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

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

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 $p(\mathcal{M}_i)$

Bayesian evidence

"Bayesian evidence" or "Marginal likelihood"

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

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:

$\lvert \ln B_{1,2} \rvert$	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)

Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

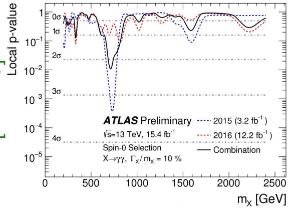
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

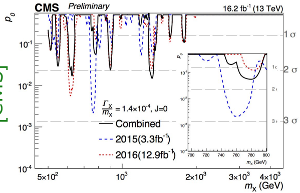
What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

750 GeV diphoton eccess

[ATLAS]



[CMS]



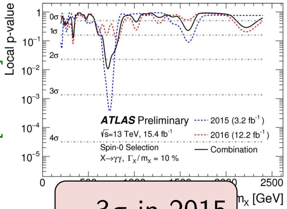
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

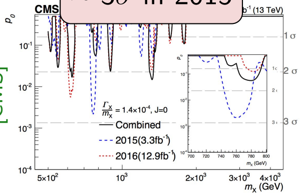
750 GeV diphoton eccess

[ATLAS]



$\sim 3\sigma$ in 2015

[CMS]



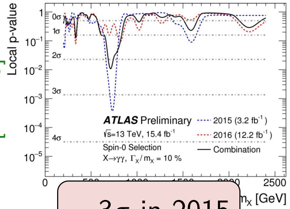
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

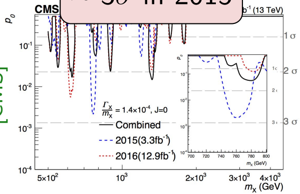
750 GeV diphoton eccess

[ATLAS]



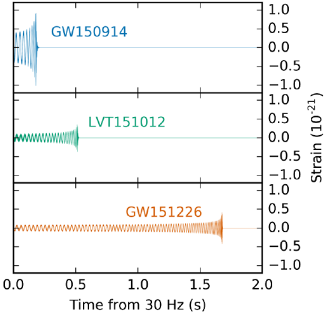
$\sim 3\sigma$ in 2015

[CMS]



LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



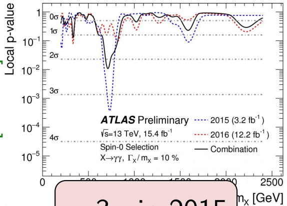
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

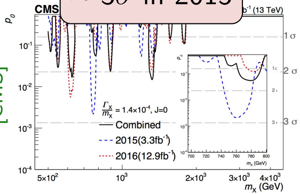
750 GeV diphoton eccess

[ATLAS]



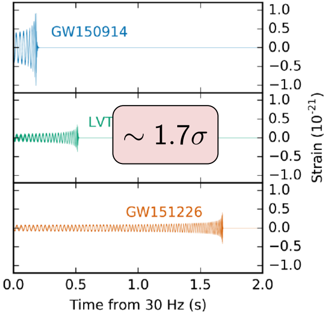
$\sim 3\sigma$ in 2015

[CMS]



LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



$\sim 1.7\sigma$

Event	$N\sigma$	$\lvert \ln B_{S+N,N} \rvert$
GW150914	$> 5.3\sigma$	
GW151226	$> 5.3\sigma$	
LVT151012	1.7σ	

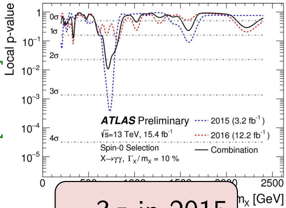
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

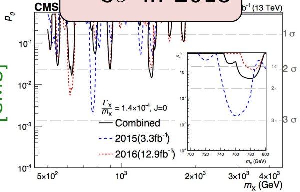
750 GeV diphoton eccess

[ATLAS]



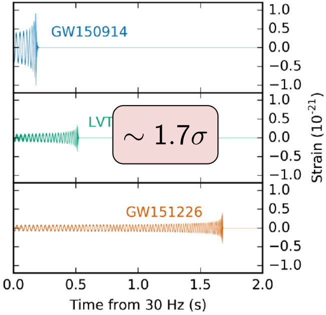
$\sim 3\sigma$ in 2015

[CMS]



LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



$\sim 1.7\sigma$

Event	$N\sigma$	$\lvert \ln B_{S+N,N} \rvert$
GW150914	$> 5.3\sigma$	~ 288
GW151226	$> 5.3\sigma$	~ 60
LVT151012	1.7σ	

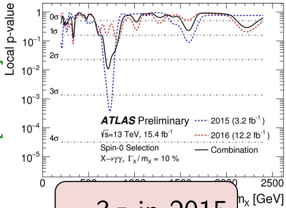
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

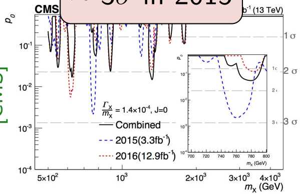
750 GeV diphoton eccess

[ATLAS]



$\sim 3\sigma$ in 2015

[CMS]



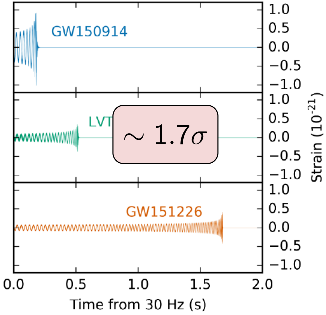
2015: $2 \lesssim \ln B_{S+N,N} \lesssim 3.7$ (prior)

2016: $\ln B_{S+N,N} \simeq -0.4$

[A.Fowlie, arxiv:1607.06608]

LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



$\sim 1.7\sigma$

Event	$N\sigma$	$\lvert\ln B_{S+N,N}\rvert$
GW150914	$> 5.3\sigma$	~ 288
GW151226	$> 5.3\sigma$	~ 60
LVT151012	1.7σ	

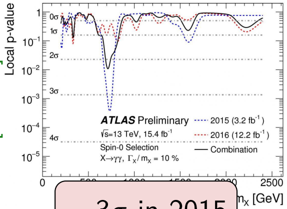
Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

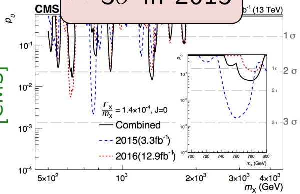
750 GeV diphoton eccess

[ATLAS]



$\sim 3\sigma$ in 2015

[CMS]



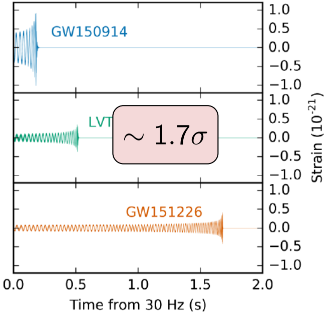
2015: $2 \lesssim \ln B_{S+N,N} \lesssim 3.7$ (prior)

2016: $\ln B_{S+N,N} \simeq -0.4$

[A.Fowlie, arxiv:1607.06608]

LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



$\sim 1.7\sigma$

Event	$N\sigma$	$\lvert \ln B_{S+N,N} \rvert$
GW150914	$> 5.3\sigma$	~ 288
GW151226	$> 5.3\sigma$	~ 60
LVT151012	1.7σ	~ 23

Frequentist significances *vs* the Bayes factor

[G. D'Agostini, arxiv:1609.01668]

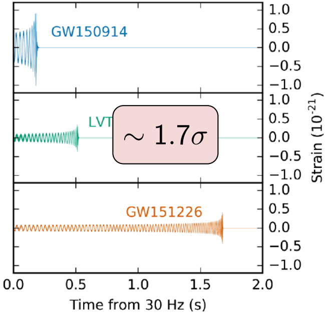
What is more robust, the frequentist "$N\sigma$" significance or the Bayes factor?

750 GeV diphoton eccess

[A. de Rujula, Bari (IT), 1995]

LIGO/Virgo trigger event

[Abbott et al., Phys.Rev.X 2016]



Event	$N\sigma$	$\lvert \ln B_{S+N,N} \rvert$
GW150914	$> 5.3\sigma$	~ 288
GW151226	$> 5.3\sigma$	~ 60
LVT151012	1.7σ	~ 23

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

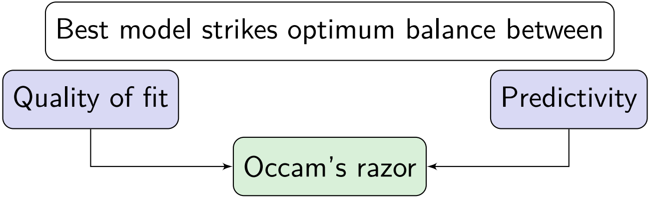
model with more parameters → better fit (usually)

→ are all the parameters needed?

Bayes factor penalizes unnecessarily complex models!

Occam's razor

what the Bayesian model comparison tells us?



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 → 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!

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

Computing the Bayesian evidence

How to compute the Bayesian evidence Z?

1. MCMCs do not work ⟶ can't explore well areas far from best-fit

2. simulated annealing

3. nested sampling [Skilling et al, 2006+]
   - `MultiNest`
   - `PolyChord`

4. approximations
   - Savage-Dickey Density Ratio (SDDR) [Dickey et al., 1970+]
   - more?

Computing the Bayesian evidence

How to compute the Bayesian evidence Z?

1. MCMCs do not work ⟶ can't explore well areas far from best-fit

2. simulated annealing [Khachaturyan et al., 1979+] ⟶ very expensive (∼100 times slower than MCMC)

3. nested sampling [Skilling et al, 2006+]
   - MultiNest
   - PolyChord

4. approximations
   - Savage-Dickey Density Ratio (SDDR) [Dickey et al., 1970+]
   - more?

Computing the Bayesian evidence

How to compute the Bayesian evidence Z?

1. MCMCs do not work ⟶ can't explore well areas far from best-fit

2. simulated annealing [Khachaturyan et al., 1979+] ⟶ very expensive (∼100 times slower than MCMC)

3. nested sampling [Skilling et al, 2006+]
   - `MultiNest` ⟶ [Feroz et al., 2008+], valid but not perfect
   - `PolyChord` ⟶ [Handley et al., 2015+], improves `MultiNest` using MCMC methods

4. approximations
   - Savage-Dickey Density Ratio (SDDR) [Dickey et al., 1970+]
   - more?

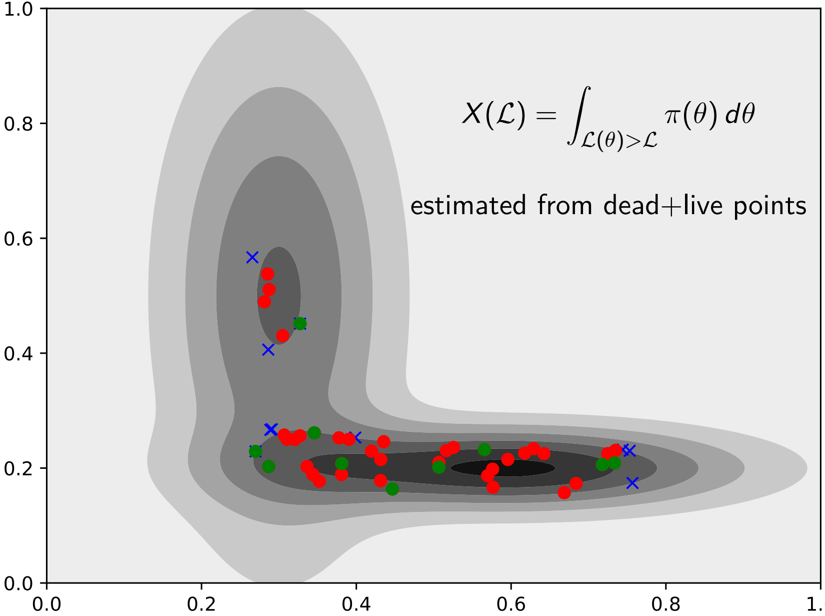
Computing the Bayesian evidence

How to compute the Bayesian evidence Z?

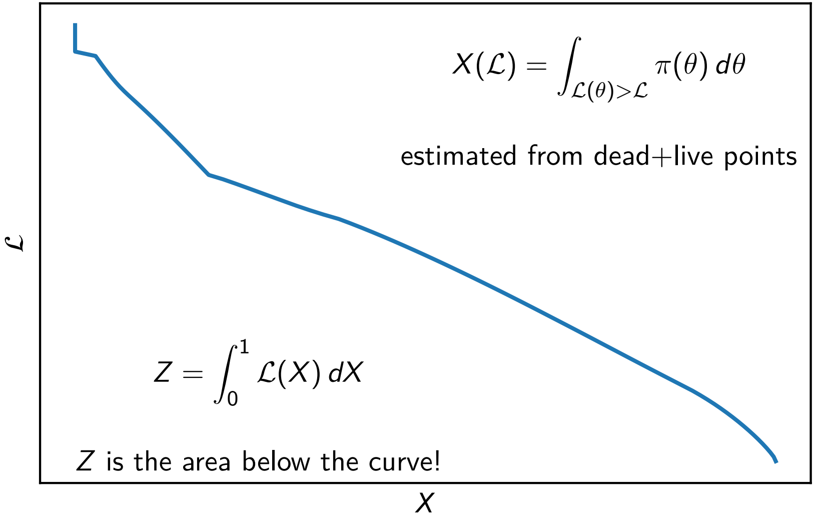
1. MCMCs do not work ⟶ can't explore well areas far from best-fit
2. simulated annealing [Khachaturyan et al., 1979+] ⟶ very expensive (∼100 times slower than MCMC)
3. nested sampling [Skilling et al, 2006+]
   - `MultiNest` ⟶ [Feroz et al., 2008+], valid but not perfect
   - `PolyChord` ⟶ [Handley et al., 2015+], improves `MultiNest` using MCMC methods
4. approximations
   - Savage-Dickey Density Ratio (SDDR) [Dickey et al., 1970+] ⟶ for nested models, $\mathcal{M}_1(\theta) \equiv \mathcal{M}_2(\theta, \psi = 0)$:

     $$B_{1,2} = \left. \frac{p(\psi|d,\mathcal{M}_2)}{p(\psi|\mathcal{M}_2)} \right|_{\psi=0}$$
   - more?

Nested sampling - a taste



Nested sampling - a taste



Neutrino masses

Normal ordering (NO)

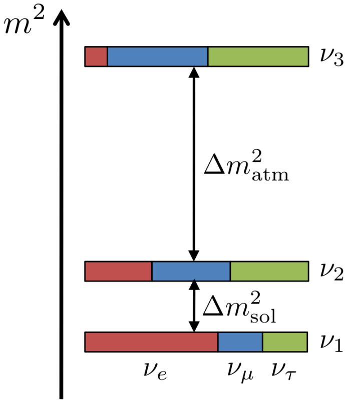
$m_1 < m_2 < m_3$

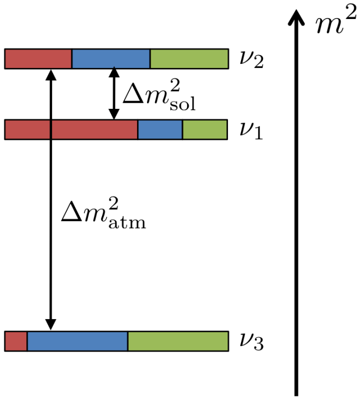
$\sum m_k \gtrsim 0.06$ eV

Inverted ordering (IO)

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV





Neutrino masses

Normal ordering (NO)

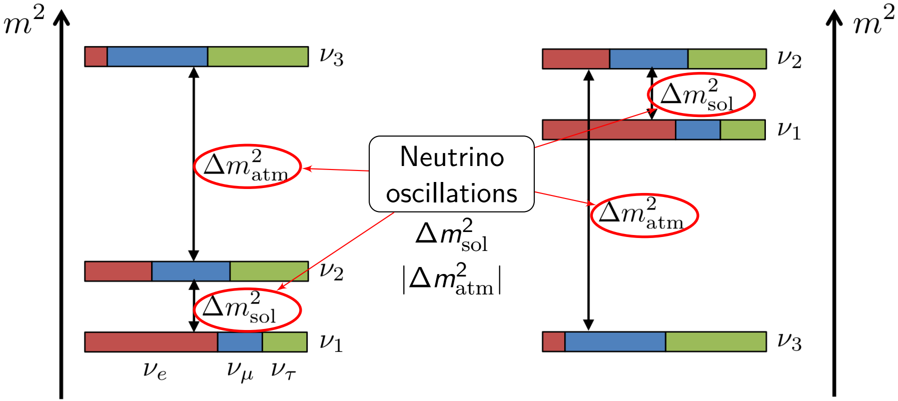
$m_1 < m_2 < m_3$

$\sum m_k \gtrsim 0.06$ eV

Inverted ordering (IO)

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV



Neutrino masses

Normal ordering (NO)

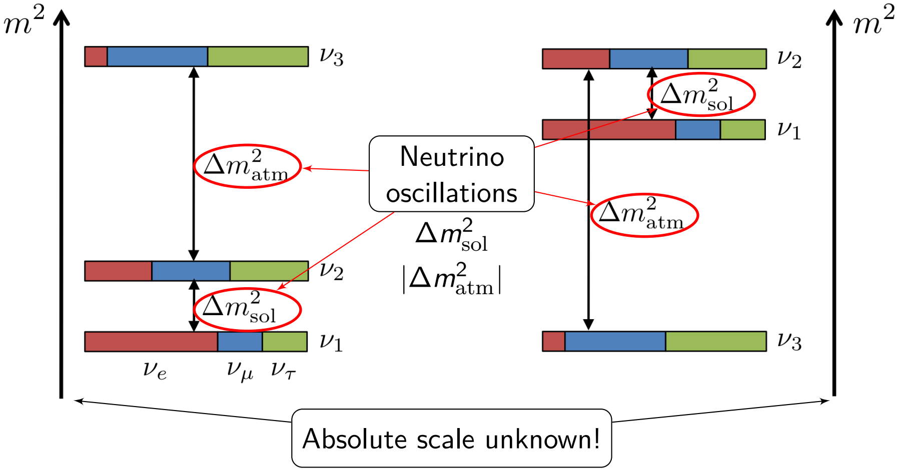
$m_1 < m_2 < m_3$

$\sum m_k \gtrsim 0.06$ eV

Inverted ordering (IO)

$m_3 < m_1 < m_2$

$\sum m_k \gtrsim 0.1$ eV



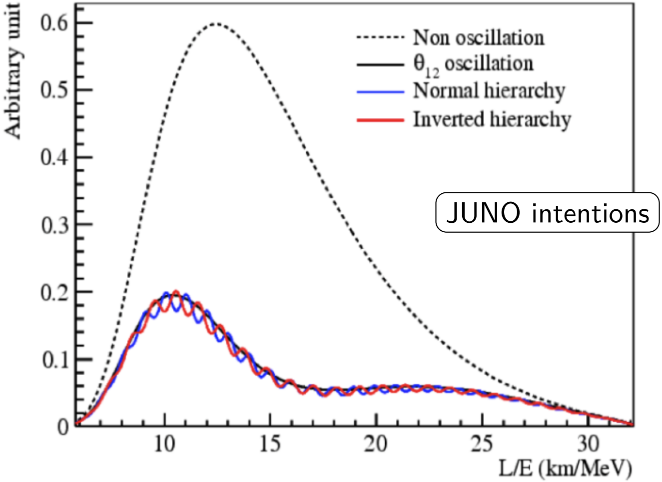
Constrain mass ordering by constraining $\sum m_k$

$\Delta m^2_{\rm atm}$ from neutrino oscillations

[de Salas et al., arxiv:1708.01186]

sign of $\Delta m^2_{\rm atm}$ from matter effects

i.e. long baseline (LBL) or atmospheric (ATM) ν experiments



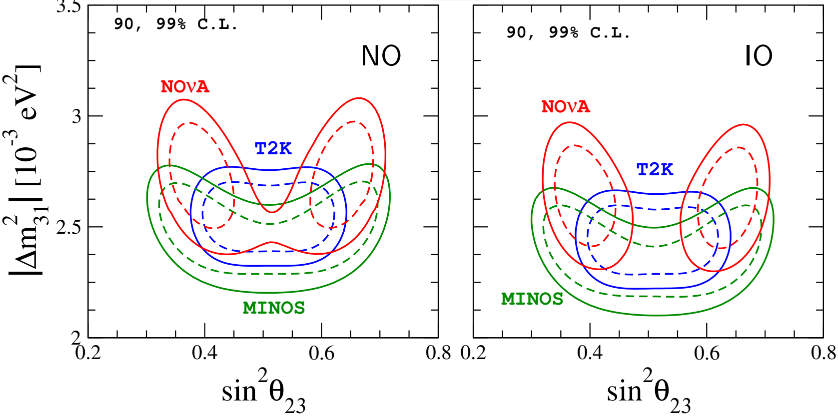
$\Delta m^2_{\rm atm}$ from neutrino oscillations

[de Salas et al., arxiv:1708.01186]

sign of $\Delta m^2_{\rm atm}$ from matter effects

i.e. long baseline (LBL) or atmospheric (ATM) ν experiments

LBL



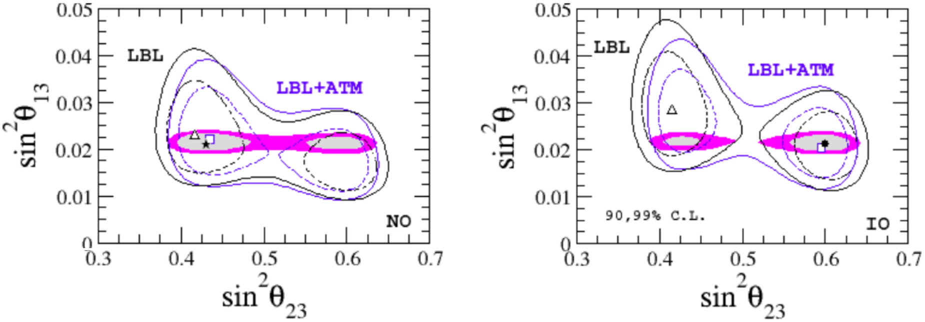
see also new T2K [M. Hartz, KEK Colloquium, Japan, 4/08/17]
and NOνA [A.Radovic, JETP seminar, Fermilab, 12/01/18] results

$\Delta m^2_{\rm atm}$ from neutrino oscillations

[de Salas et al., arxiv:1708.01186]

sign of $\Delta m^2_{\rm atm}$ from matter effects

i.e. long baseline (LBL) or atmospheric (ATM) ν experiments



Neutrino masses from β decay

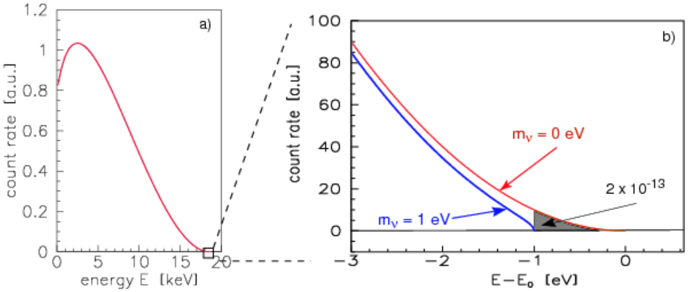
Must measure β decay endpoint → $m_{\nu_e}^2 = \sum_k |U_{ek}|^2 m_k^2$

U_{ek} mixing matrix

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

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

[Katrin L.o.I., 2001]



Neutrino masses from β decay

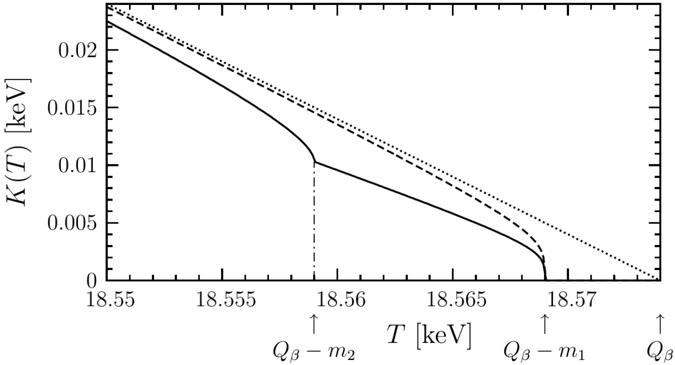
Must measure β decay endpoint → $m^2_{\nu_e} = \sum_k |U_{ek}|^2 \, m^2_k$

U_{ek} mixing matrix

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

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

[Giunti&Kim, 2007]



Neutrino masses from neutrinoless double β decay

(if neutrino is Majorana)

[Schechter&Valle, 1982]

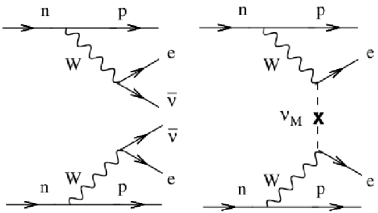
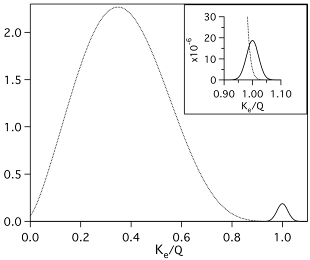
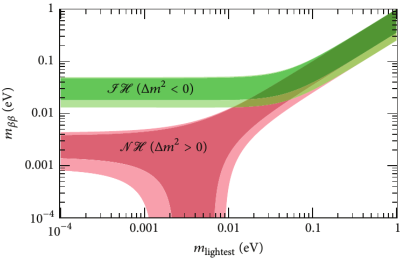


figure from [NEXT] webpage





[Dell'Oro et al., 2016]

Measure $T^{0\nu}_{1/2}$

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

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

α_k Majorana phases

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

Neutrino masses from CMB

ω_i energy density of species i,
$i \in$ (radiation, matter, baryons, cold dark matter, ν)
$z_{\rm eq}$ matter-radiation equality redshift

$$1 + z_{\rm eq} = (\omega_b + \omega_c)/\omega_r$$

independent of m_ν

$$\omega_m^0 = \omega_b^0 + \omega_c^0 + \omega_\nu^0 \text{ today}$$

mass of species relativistic at recombination
affects late time evolution only

small effects on the SW plateau
(cosmic variance, degeneracies...)

Effects on the early ISW effect

$$\frac{\Delta C_\ell}{C_\ell} \simeq -\left(\frac{\sum m_\nu}{0.1 \text{ eV}}\right)\%$$

effects on the position of peaks

$$\theta_s = r_s(\eta_{LS})/D_A(\eta_{LS})$$

$$D_A = \int_0^{z_{rec}} \frac{dz}{H(z)}$$

(this effect can be compensated reducing H_0)

correlation m_ν–H_0

["Neutrino Cosmology", Lesgourgues et al.]

Neutrino masses from CMB

ω_i energy density of species i,
$i \in$ (radiation, matter, baryons, cold dark matter, ν)
$z_{\rm eq}$ matter-radiation equality redshift

$$1 + z_{\rm eq} = (\omega_b + \omega_c)/\omega_r$$

independent of m_ν

$$\omega_m^0 = \omega_b^0 + \omega_c^0 + \omega_\nu^0 \text{ today}$$

mass of species relativistic at recombination affects late time evolution only

small effects on the SW plateau (cosmic variance, degeneracies...)

Effects on the early ISW effect

$$\frac{\Delta C_\ell}{C_\ell} \simeq -\left(\frac{\sum m_\nu}{0.1 \text{ eV}}\right)\%$$

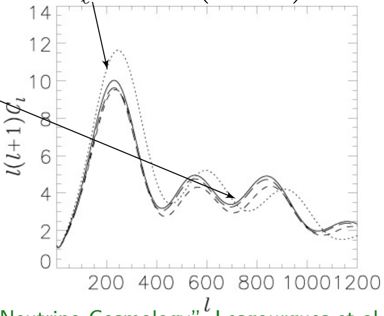
effects on the position of peaks

$$\theta_s = r_s(\eta_{LS})/D_A(\eta_{LS})$$

$$D_A = \int_0^{z_{rec}} \frac{dz}{H(z)}$$

(this effect can be compensated reducing H_0)

correlation m_ν–H_0



["Neutrino Cosmology", Lesgourgues et al.]

Cosmological mass bounds

Bounds on $M_\nu = \sum m_\nu$

standard

based on ΛCDM model

[Planck Collaboration 2015, AA594 (2016) A13]

95%

$M_\nu < 0.72$ eV (PlanckTT+lowP)

$M_\nu < 0.21$ eV (+BAO)

$M_\nu < 0.49$ eV (PlanckTTTEEE+lowP)

$M_\nu < 0.17$ eV (+BAO)

see also:

[Vagnozzi et al., PRD96 (2017) 123503]

[Planck Collaboration 2016, AA596 (2016) A107]

95%

$M_\nu < 0.59$ eV (PlanckTT+SimLow)

$M_\nu < 0.17$ eV (+BAO)

$M_\nu < 0.34$ eV (PlanckTTTEEE+SimLow)

$M_\nu < 0.14$ eV (+BAO)

(SimLow not public yet)

Cosmological mass bounds

Bounds on $M_\nu = \sum m_\nu$

standard

based on ΛCDM model

[Planck Collaboration 2015, AA594 (2016) A13]

95%
- $M_\nu < 0.72$ eV (PlanckTT+lowP)
- $M_\nu < 0.21$ eV (+BAO)
- $M_\nu < 0.49$ eV (PlanckTTTEEE+lowP)
- $M_\nu < 0.17$ eV (+BAO)

see also:

[Vagnozzi et al., PRD96 (2017) 123503]

[Planck Collaboration 2016, AA596 (2016) A107]

95%
- $M_\nu < 0.59$ eV (PlanckTT+SimLow)
- $M_\nu < 0.17$ eV (+BAO)
- $M_\nu < 0.34$ eV (PlanckTTTEEE+SimLow)
- $M_\nu < 0.14$ eV (+BAO)

(SimLow not public yet)

Modified gravity?

[Barreira et al., 2014]:

νGalileon

68%
- $M_\nu = 0.98 \pm 0.24$ eV (CMB)
- $M_\nu = 0.65 \pm 0.11$ eV (CMB+BAO)

[Bellomo et al., 2016]:

95% Horndeski scalar-tensor

$M_\nu < 0.76$ eV

[Dirian, 2017]:

68% nonlocal gravity

$M_\nu = 0.21 \pm 0.08$ eV

[Peirone et al, 2017]:

68% Covariant Galileon

$M_\nu = 0.8 \pm 0.1$ eV

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

Can current data tell us the neutrino mass ordering?

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

NEUTRINO ORDERING WARS

Episode VIII

THE LAST PRIOR

arXiv:1801.04946v1 [hep-ph]

Neutrino masses and their ordering: Global Data, Priors and Models

S. Gariazzo,[a] M. Archidiacono,[b] P.F. de Salas,[a] O. Mena,[a] C.A. Ternes,[a] and M. Tórtola[a]

[a]Instituto de Física Corpuscular (CSIC-Universitat de València)
Parc Científic UV, C/ Catedrático José Beltrán, 2
E-46980 Paterna (Valencia), Spain

[b]Institute for Theoretical Particle Physics and Cosmology (TTK)
RWTH Aachen University, D-52056 Aachen, Germany

Parameterizations, priors and data

[Gariazzo et al., arxiv:1801.04946]

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., arxiv:1801.04946]

$0\nu\beta\beta$ 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]

$0\nu\beta\beta$		Neutrino mixing	
Parameter	Prior	Parameter	Prior
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		

Masses: see later!

Parameterizations, priors and data

[Gariazzo et al., arxiv:1801.04946]

Cosmological data

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

$0\nu\beta\beta$ 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		$0\nu\beta\beta$		Neutrino mixing	
Parameter	Prior	Parameter	Prior	Parameter	Prior
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				

Masses: see later!

Parameterizing neutrino masses

[Gariazzo et al., arxiv:1801.04946]

[Simpson et al, 2017]

using m_1, m_2, m_3 (A)

[Caldwell et al, 2017]

using $m_{\rm lightest}$, Δm^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?

Parameterizing neutrino masses

[Gariazzo et al., arxiv:1801.04946]

[Simpson et al, 2017]

using m_1, m_2, m_3 (A)

[Caldwell et al, 2017]

using $m_{\rm lightest}$, Δm^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 ($m_{\rm lightest}$)?

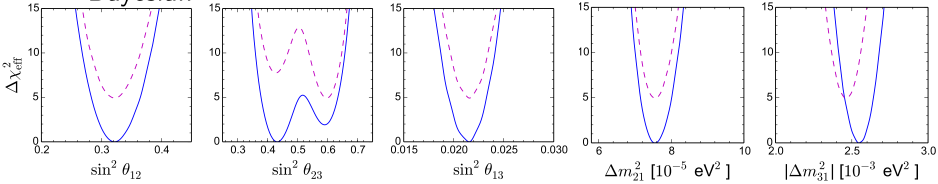
Can data help to select (A) or (B), linear or log?

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

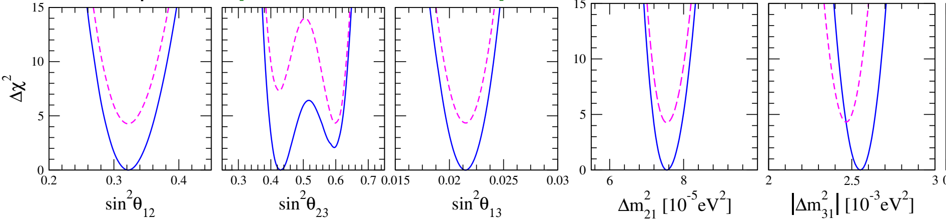
Neutrino mixing results

[Gariazzo et al., arxiv:1801.04946]

Bayesian



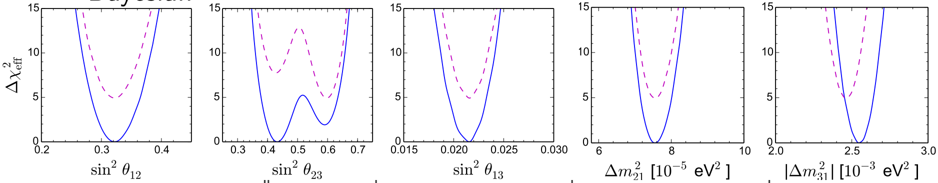
Frequentist [de Salas et al., 1708.01186]



Neutrino mixing results

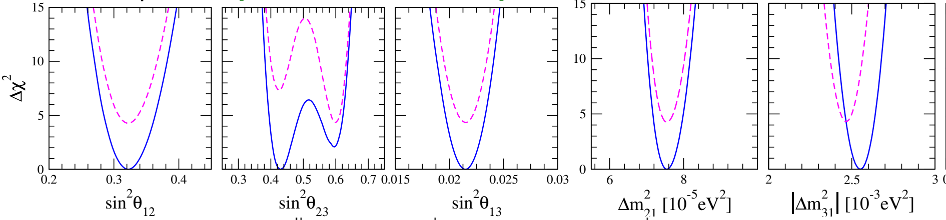
[Gariazzo et al., arxiv:1801.04946]

Bayesian



parameter	bestfit	1σ range	2σ range	3σ range
$\sin^2\theta_{23}$ (NO)	0.43 (0.59)	0.40–0.47 & 0.58–0.60	0.38–0.50 & 0.53–0.63	0.37–0.65
$\sin^2\theta_{23}$ (IO)	0.60 (0.43)	0.56–0.63	0.39–0.47 & 0.54–0.65	0.37–0.66

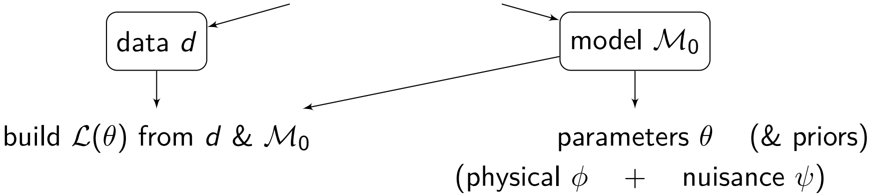
Frequentist [de Salas et al., 1708.01186]



parameter	best fit $\pm 1\sigma$	2σ range	3σ range
$\sin^2\theta_{23}/10^{-1}$ (NO)	$4.30^{+0.20}_{-0.18}$ [a]	3.98–4.78 & 5.60–6.17	3.84–6.35
$\sin^2\theta_{23}/10^{-1}$ (IO)	$5.98^{+0.17}_{-0.15}$ [b]	4.09–4.42 & 5.61–6.27	3.89–4.88 & 5.22–6.41

(Bayesian) Parameter inference

Parameter inference = what we learn on the parameters, given:



Full posterior:

$$p(\theta|d,\mathcal{M}_0) \propto \mathcal{L}(\theta) \times p(\theta|\mathcal{M}_0)$$

(Bayesian) Parameter inference

Parameter inference = what we learn on the parameters, given:

data d → build $\mathcal{L}(\theta)$ from d & $\mathcal{M}_0$

model $\mathcal{M}_0$ → parameters θ (& priors)

(physical ϕ + nuisance ψ)

Full posterior:

$$p(\theta|d, \mathcal{M}_0) \propto \mathcal{L}(\theta) \times p(\theta|\mathcal{M}_0)$$

Marginalize over nuisance to obtain posterior for physical:

$$p(\phi|d, \mathcal{M}_0) \propto \int \mathcal{L}(\phi, \psi) p(\phi, \psi|\mathcal{M}_0) d\psi$$

marginalize over all the parameters except one (two)

→ 1D (2D) posterior

Credible intervals from the posterior

Credible interval α?

range of values within which an unobserved parameter value falls with a particular subjective probability α

Credible intervals from the posterior

Credible interval α?

range of values within which an unobserved parameter value falls
with a particular subjective probability α

Analogous to frequentist confidence intervals α

Bayesian credible interval:

- bounds as fixed;
- estimated parameter as a random variable.

Frequentist confidence interval:

- bounds as random variables;
- estimated parameter as fixed value.

Credible intervals from the posterior

Credible interval α?

range of values within which an unobserved parameter value falls with a particular subjective probability α

Analogous to frequentist confidence intervals α

Bayesian credible interval:

- bounds as fixed;
- estimated parameter as a random variable.

Frequentist confidence interval:

- bounds as random variables;
- estimated parameter as fixed value.

Credible intervals are not uniquely defined!

highest posterior density interval: narrowest interval, includes values of highest probability density

equal-tailed interval: same probability of being below or above the interval

interval for which the mean is the central point

Credible intervals from the posterior

Credible interval α?

range of values within which an unobserved parameter value falls with a particular subjective probability α

Analogous to frequentist confidence intervals α

Bayesian credible interval:

- bounds as fixed;
- estimated parameter as a random variable.

Frequentist confidence interval:

- bounds as random variables;
- estimated parameter as fixed value.

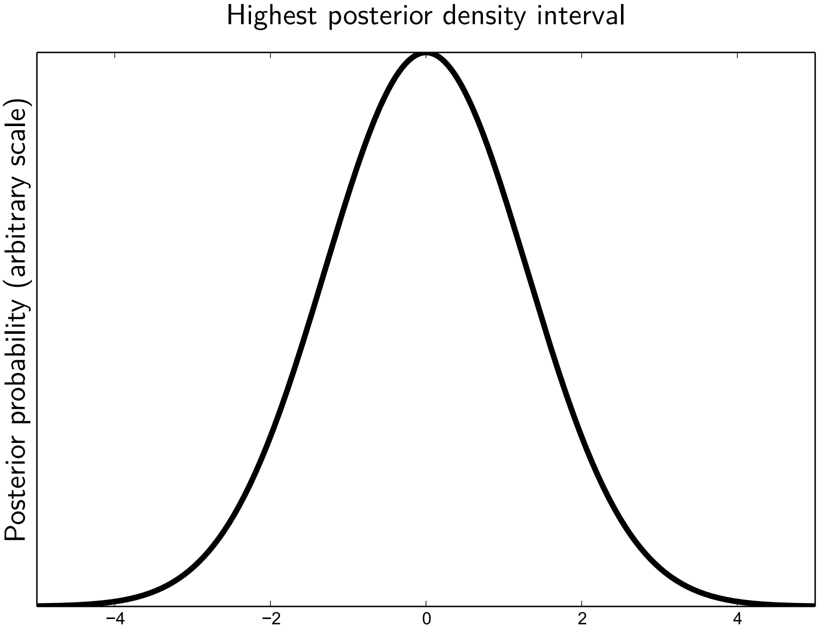
Credible intervals are not uniquely defined!

highest posterior density interval: narrowest interval, includes values of highest probability density

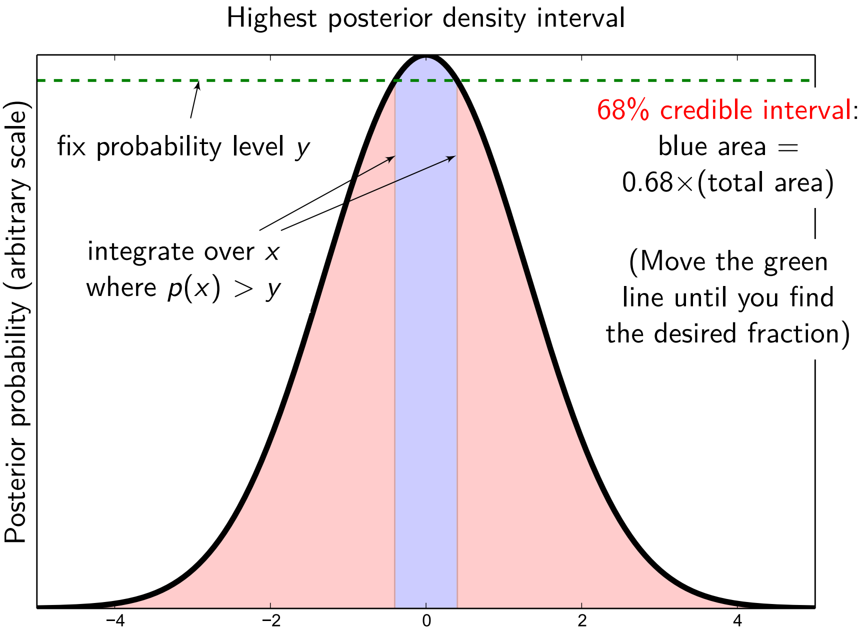
equal-tailed interval: same probability of being below or above the interval

interval for which the mean is the central point

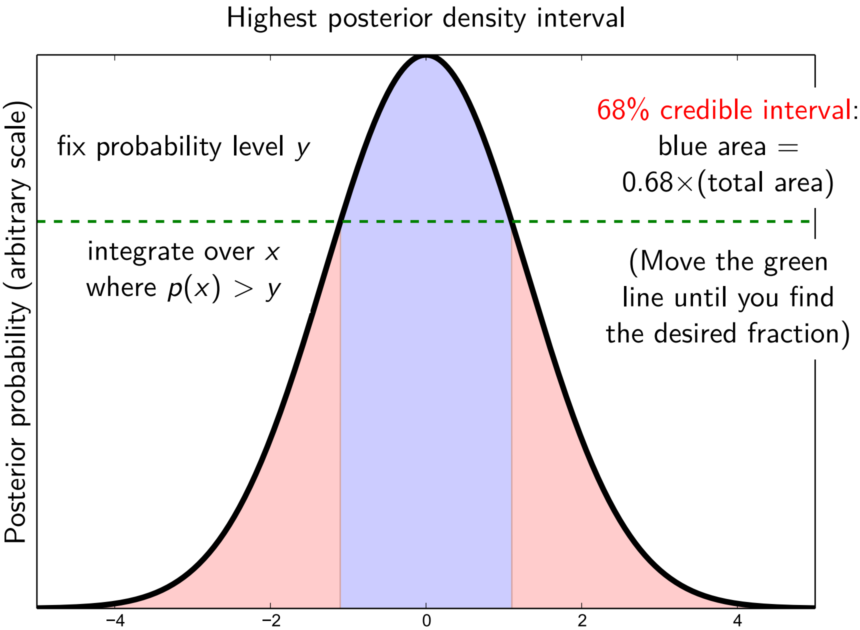
Computing credible intervals



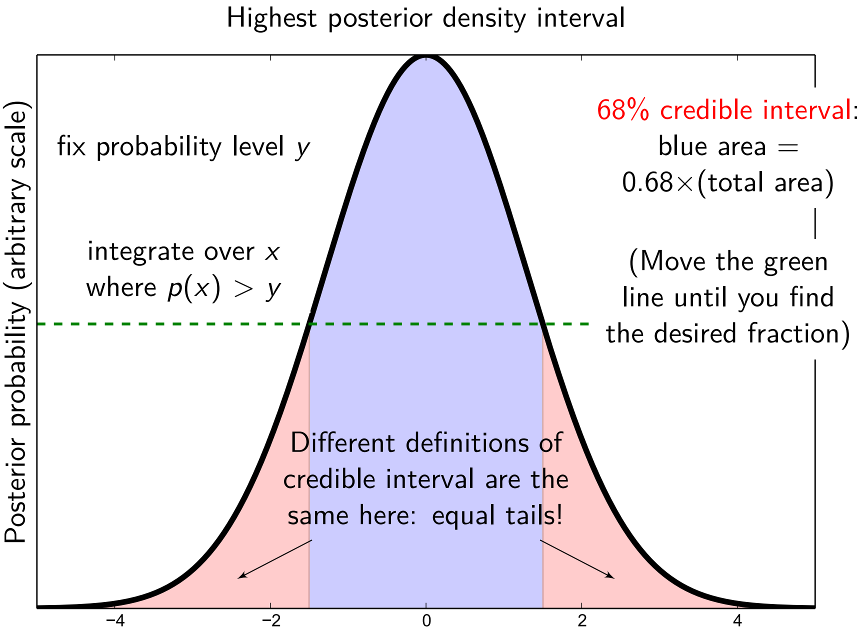
Computing credible intervals



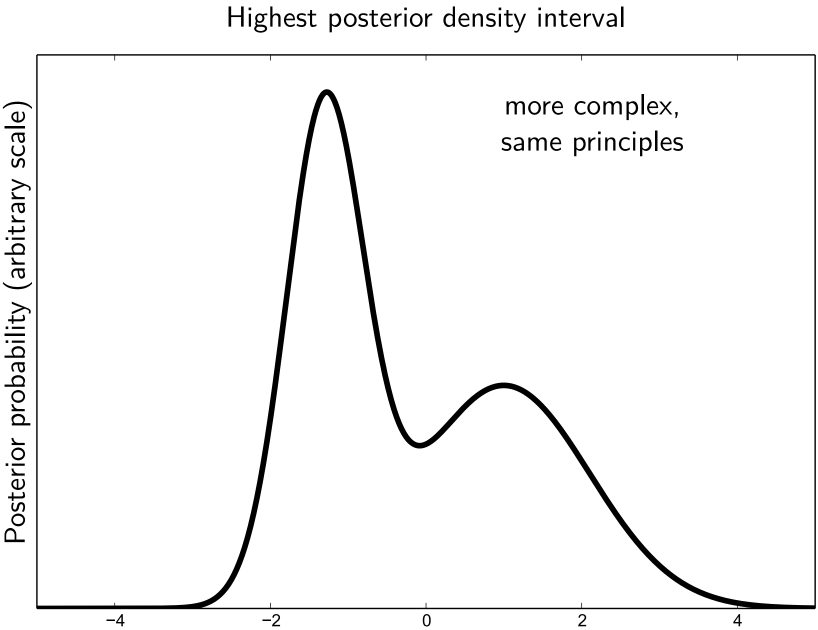
Computing credible intervals



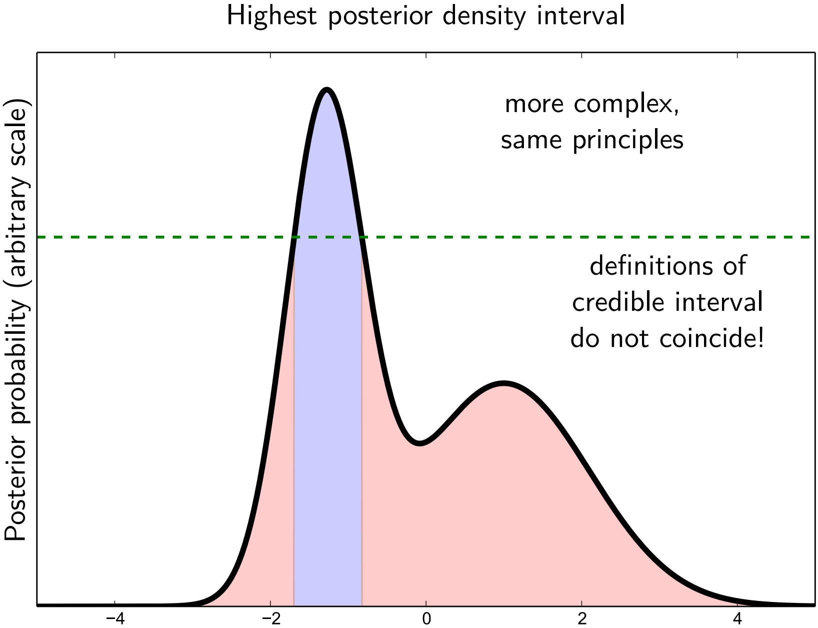
Computing credible intervals



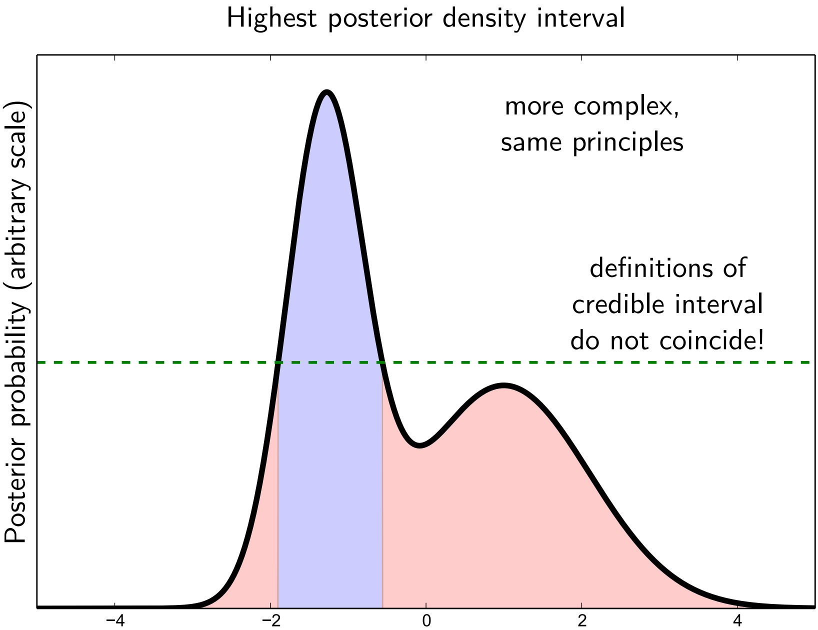
Computing credible intervals



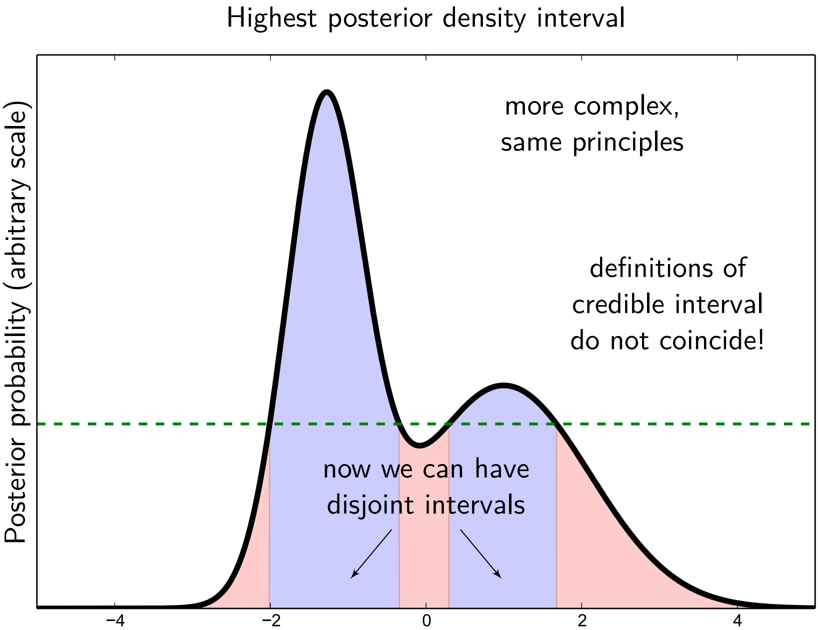
Computing credible intervals



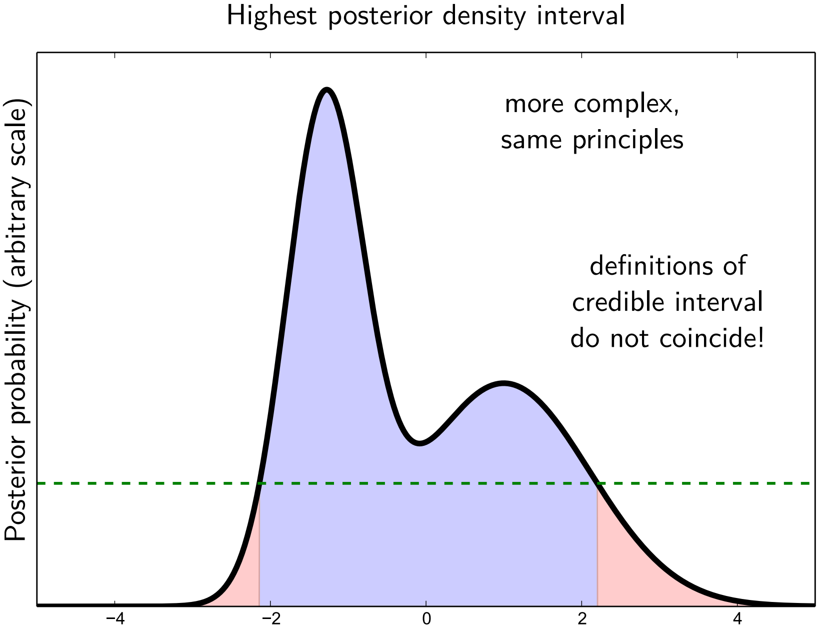
Computing credible intervals



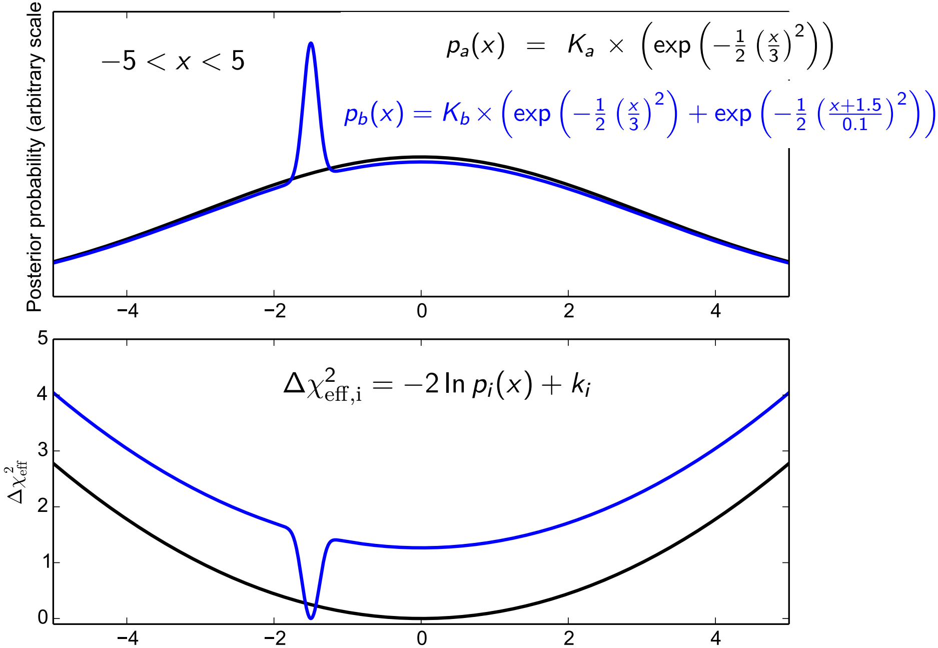
Computing credible intervals



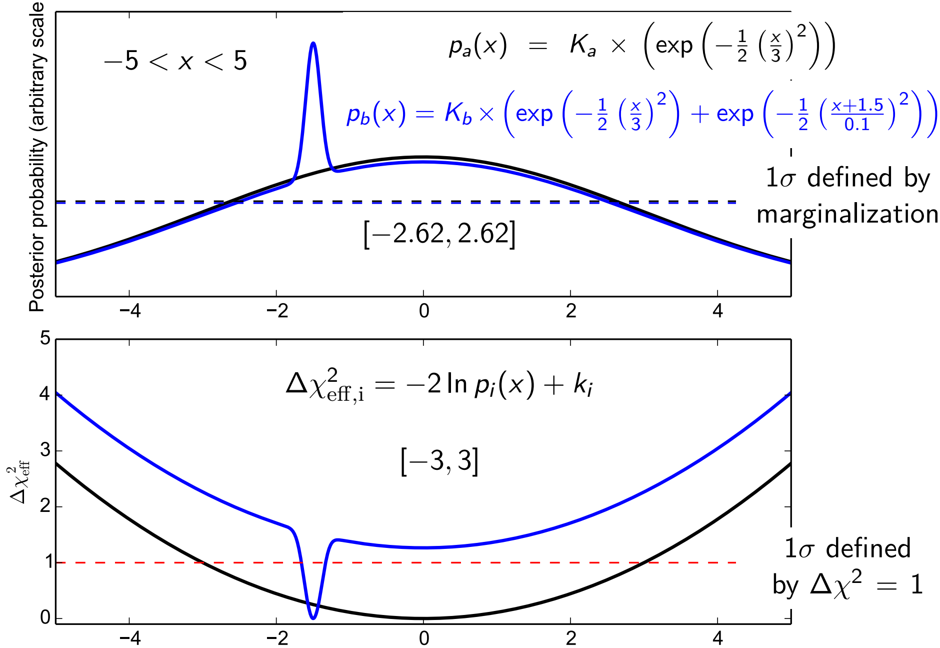
Computing credible intervals



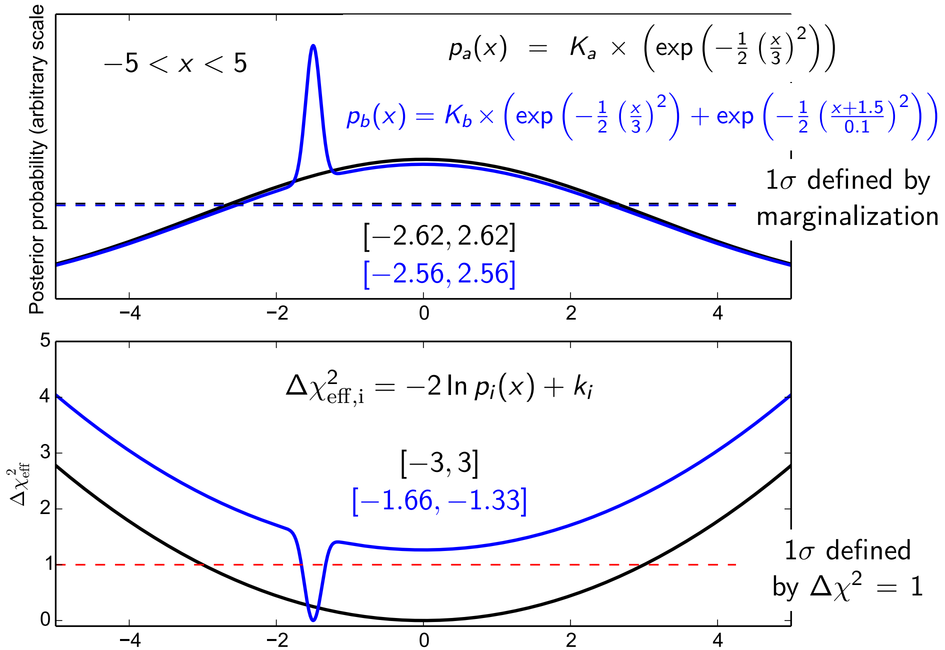
Where Bayesian and frequentist results differ



Where Bayesian and frequentist results differ

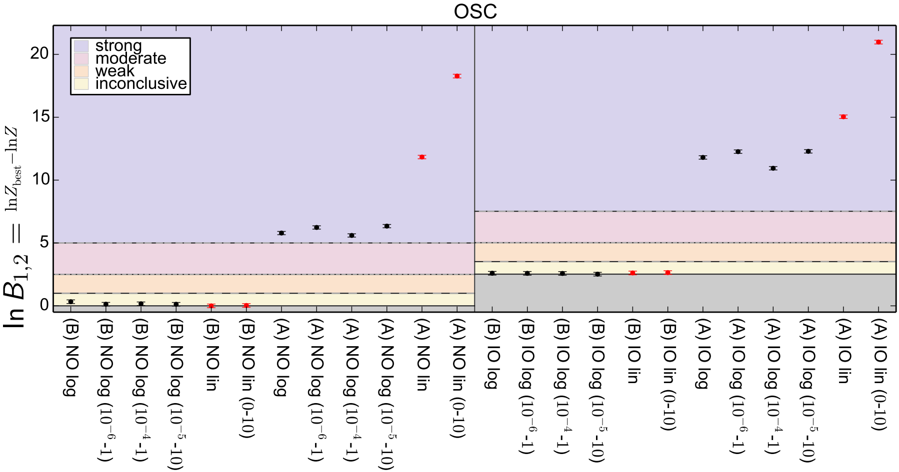


Where Bayesian and frequentist results differ



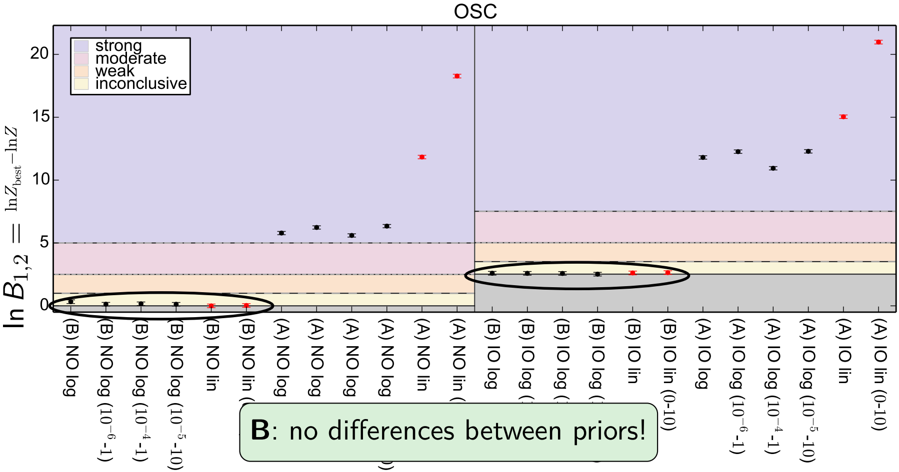
Comparing parameterizations/priors

[Gariazzo et al., arxiv:1801.04946]



Comparing parameterizations/priors

[Gariazzo et al., arxiv:1801.04946]

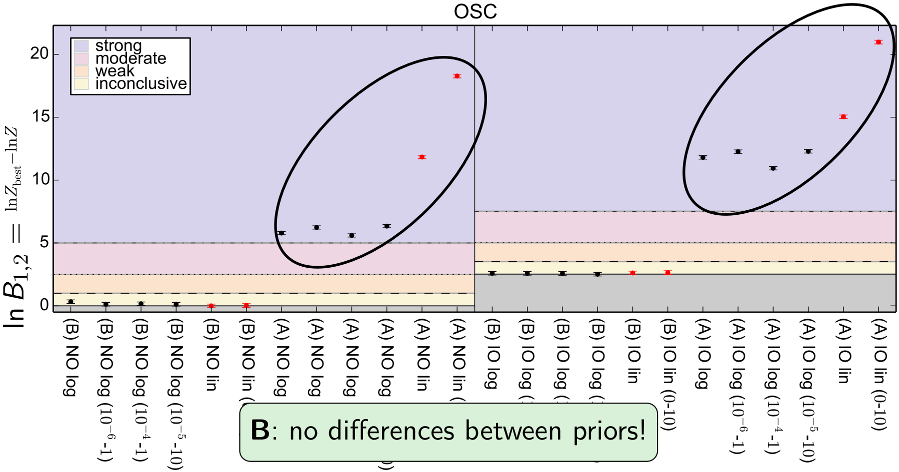


B: no differences between priors!

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

Comparing parameterizations/priors

[Gariazzo et al., arxiv:1801.04946]



B: no differences between priors!

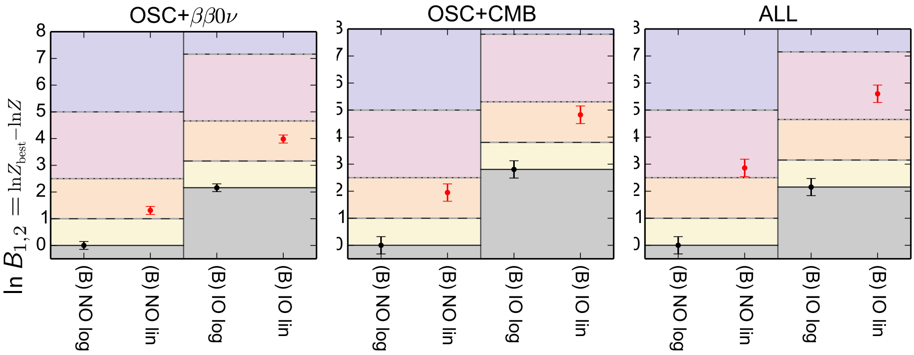
(no penalty for unconstrained parameter $m_{\rm lightest}$!)

A: always strongly disfavored!

(waste of parameter space, no unconstrained parameters due to Δm^2_{i1}!)

Comparing parameterizations/priors

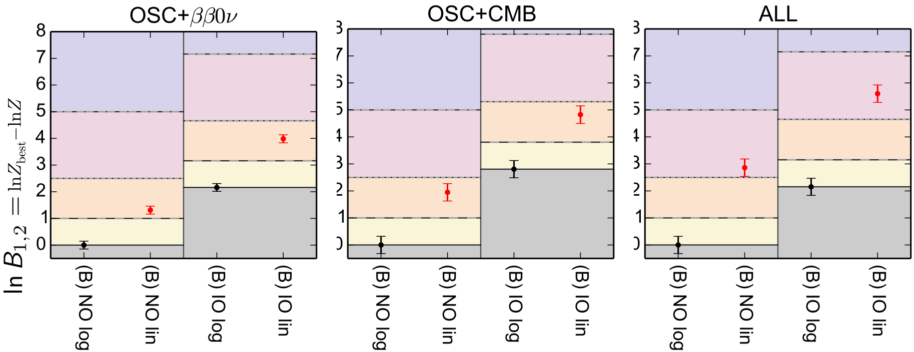
[Gariazzo et al., arxiv:1801.04946]



compare linear versus **logarithmic**

Comparing parameterizations/priors

[Gariazzo et al., arxiv:1801.04946]

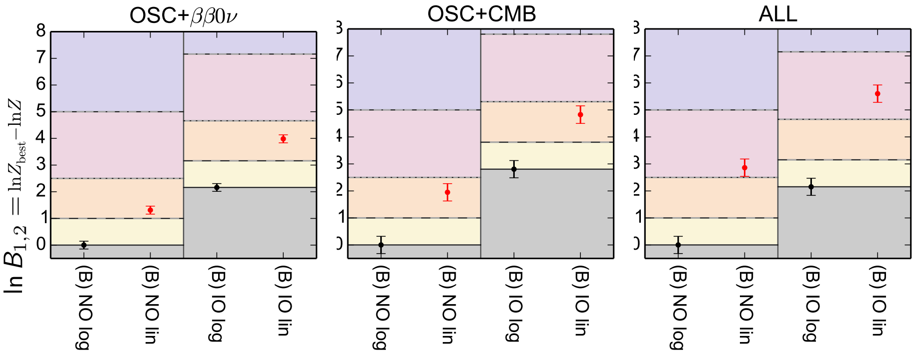


compare linear versus **logarithmic**

log priors are
weakly-to-moderately more efficient

Comparing parameterizations/priors

[Gariazzo et al., arxiv:1801.04946]



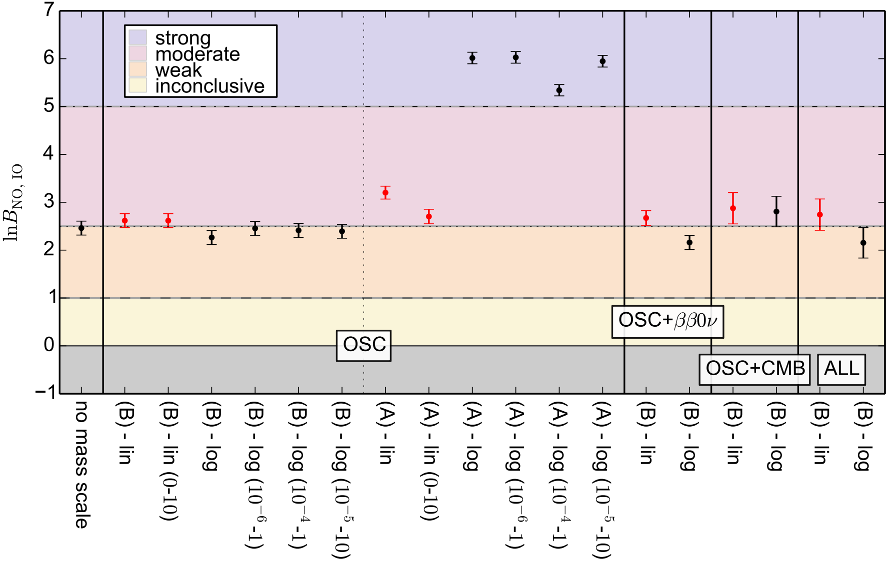
compare linear versus **logarithmic**

log priors are
weakly-to-moderately more efficient

summary: model B, log prior is better!

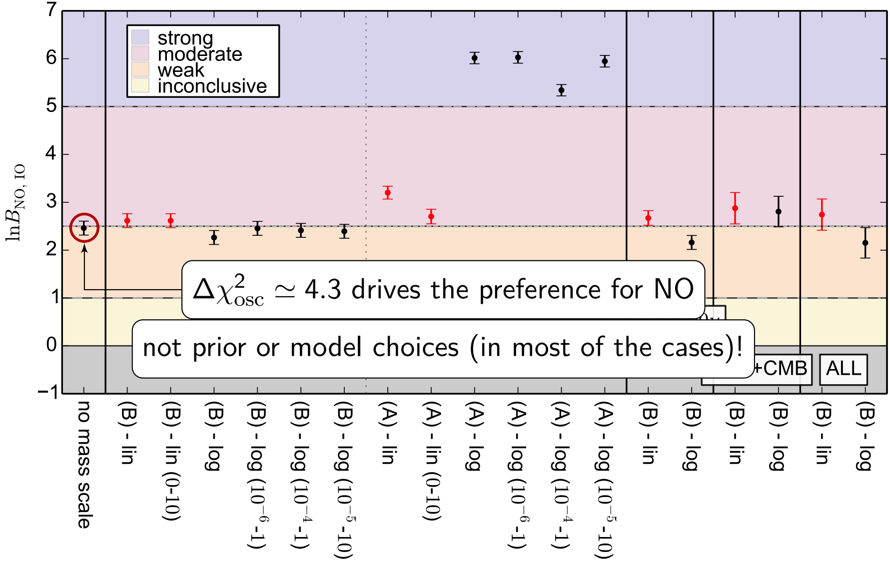
Comparing the mass orderings

[Gariazzo et al., arxiv:1801.04946]



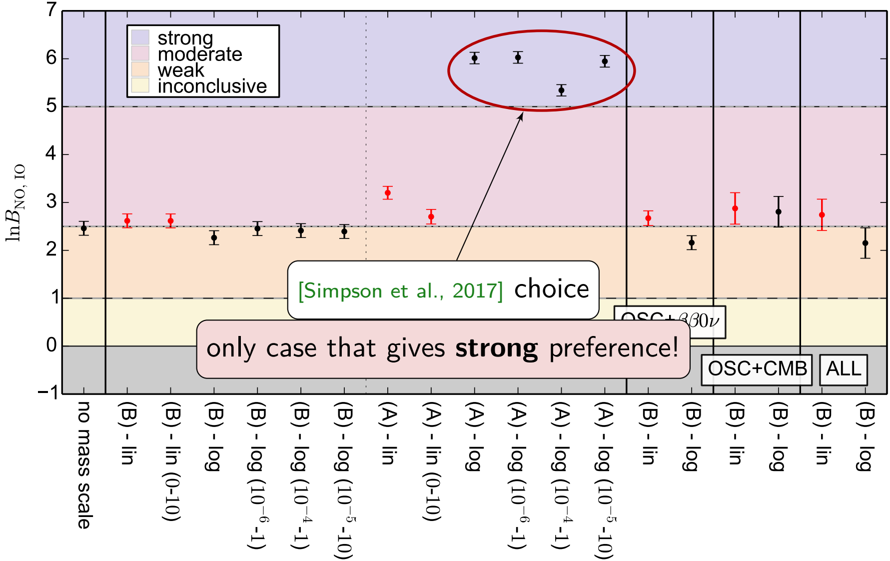
Comparing the mass orderings

[Gariazzo et al., arxiv:1801.04946]



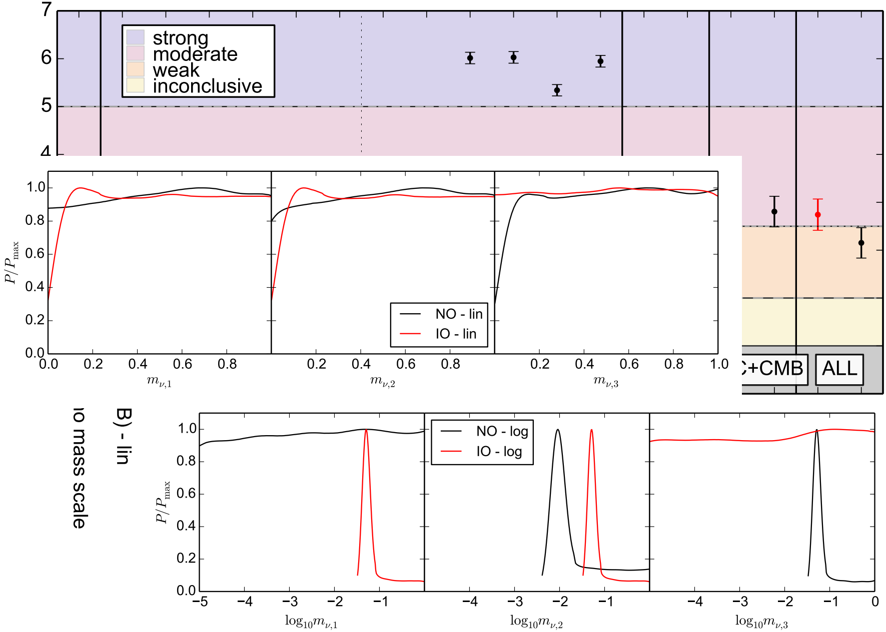
Comparing the mass orderings

[Gariazzo et al., arxiv:1801.04946]



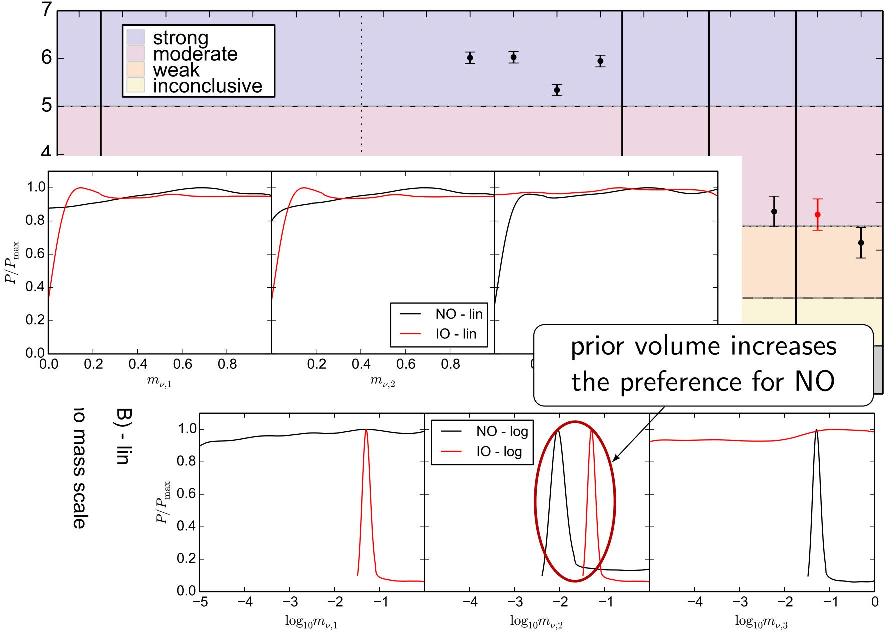
Comparing the mass orderings

[Gariazzo et al., arxiv:1801.04946]



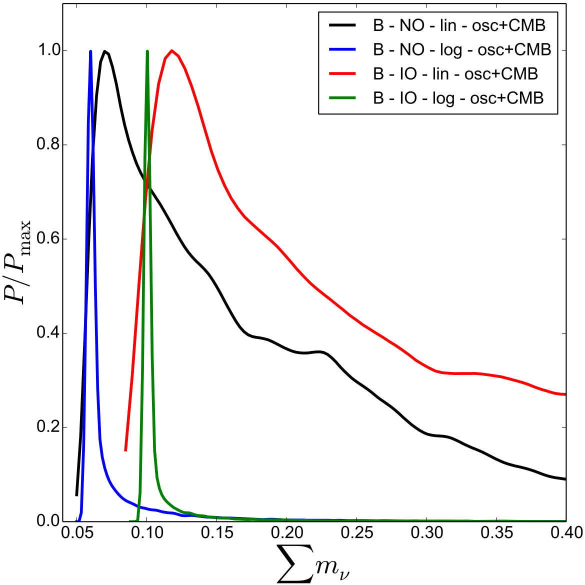
Comparing the mass orderings

[Gariazzo et al., arxiv:1801.04946]



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

[Gariazzo et al., arxiv:1801.04946]

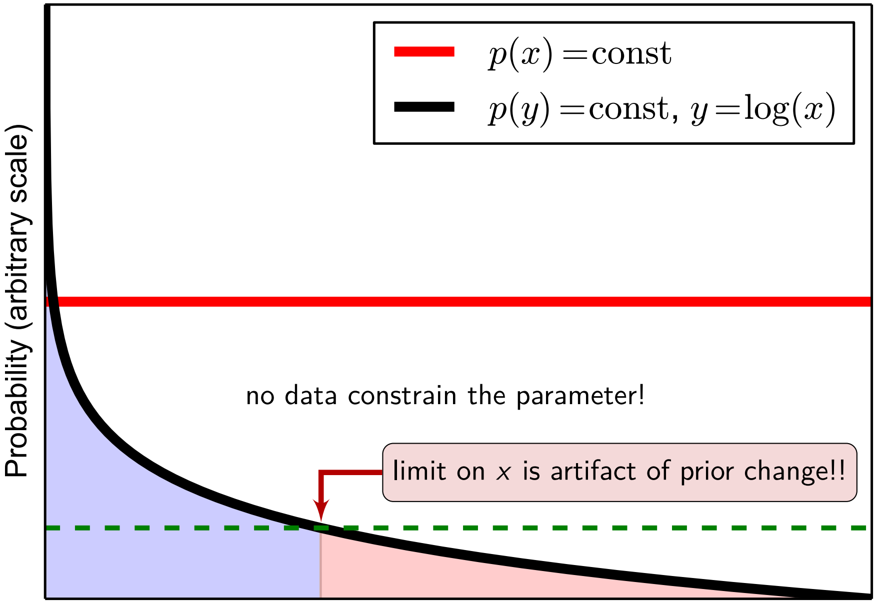


showing model B (1 mass parameter)

would be the same for model A, but amplified (3 mass parameters!)

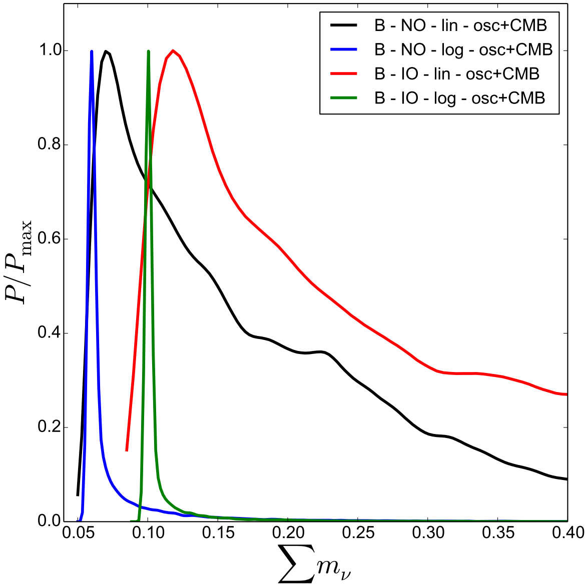
The role of priors: $\sum m_\nu$

[Gariazzo et al., arxiv:1801.04946]



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

[Gariazzo et al., arxiv:1801.04946]



showing model B (1 mass parameter)

would be the same for model A, but amplified (3 mass parameters!)

logarithmic prior corresponds to $1/m_k$ probability!

↓

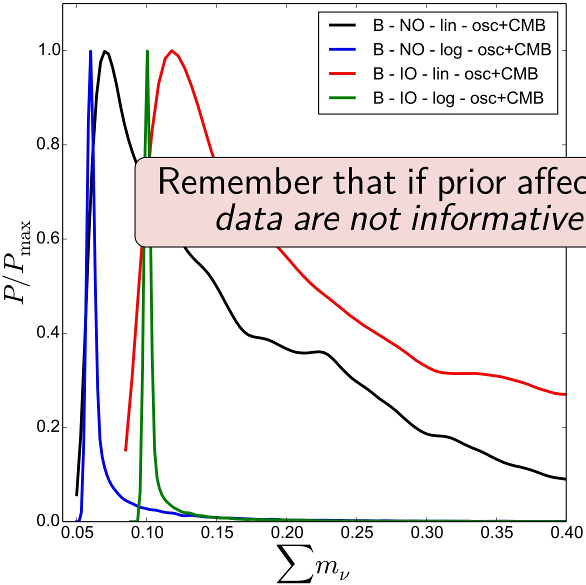
more importance to smaller masses

↓

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

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

[Gariazzo et al., arxiv:1801.04946]



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

showing model B
(1 mass parameter)

would be the same
for model A, but
3 mass
ers!)

logarithmic prior
corresponds to
$1/m_k$ probability!

↓

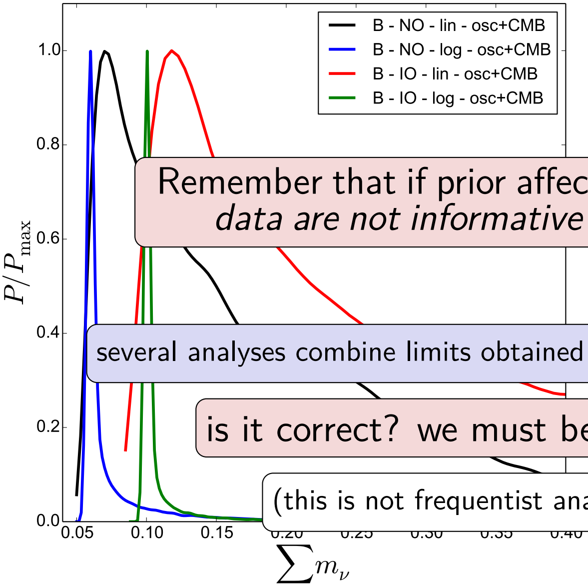
more importance
to smaller masses

↓

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

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

[Gariazzo et al., arxiv:1801.04946]



showing model B (1 mass parameter)

would be the same for model A, but ... 3 mass ... ers!)

logarithmic prior corresponds to ... ty!

↓ ... importance to smaller 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? we must be careful!

(this is not frequentist analysis)

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
and when combining different analyses

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
and when combining different analyses

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

Thank you for the attention!