

Bayesian generalised mixed models

with *MCMCglmm*

Ferran Sayol



@FerranSayol



fsayol@gmail.com



Chocolate Hills, Bohol, Philippines

Bayesian generalised **mixed models**

with *MCMCglmm*

Bayesian generalised mixed models

with *MCMCglmm*

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$$

Bayesian generalised mixed models

with *MCMCglmm*

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$$

A **mixed model** is a statistical model containing both **fixed** effects and **random** effects.

Fixed effects

Our predictors. Some specific trait of the samples that explain what we measure.

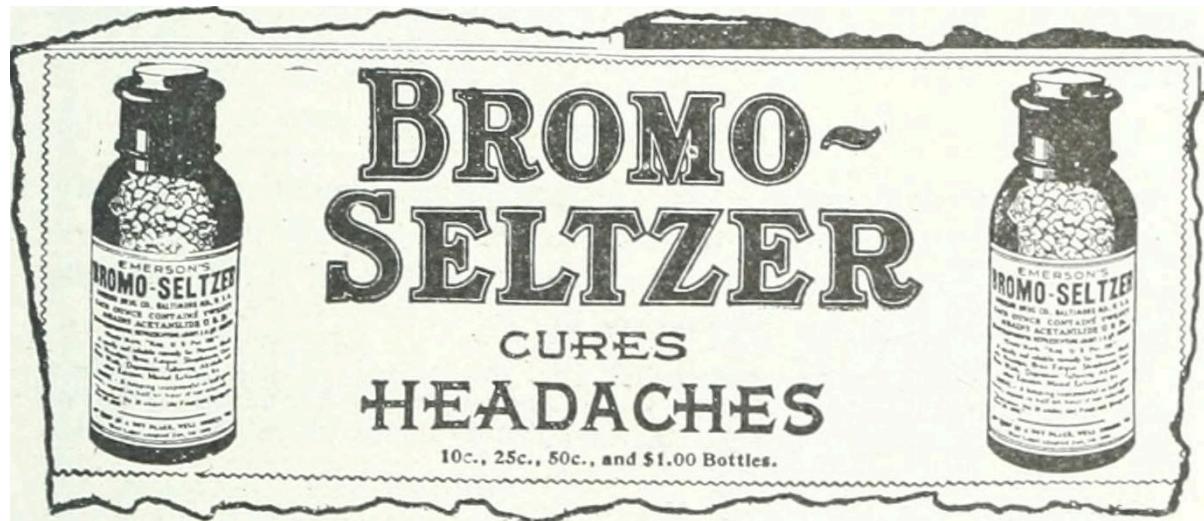
Data has been gathered from all the levels of the factor that are of interest.

Random effects

Some arbitrary "label" of our measure that add extra variation.

The factor has many possible levels, but only a random sample of levels is included in the data.

Testing a new drug



Fixed effects

Age, sex, pregnancy.

Random effects

Hospital, City, Person.

In a way, we want to control the variation due to uncontrolled (~random) noise, so we can see if the fixed terms have significant effects.

Bayesian generalised mixed models with *MCMCglmm*

Bayesian generalised mixed models

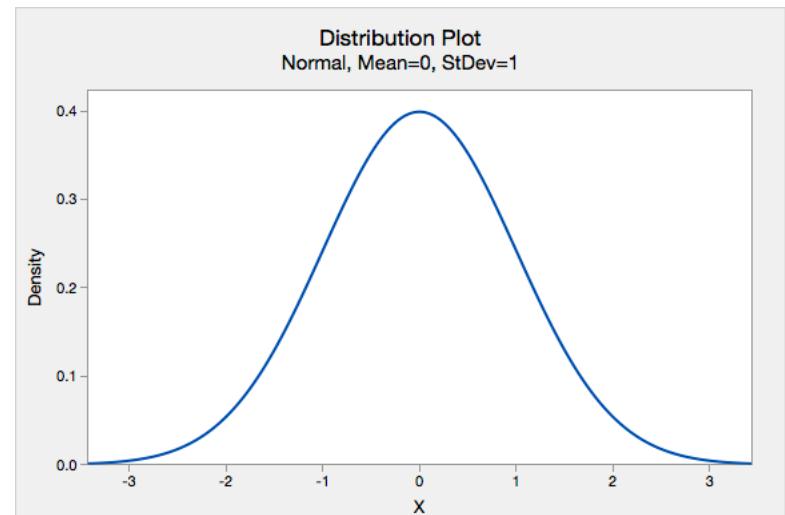
with *MCMCglmm*

Normal distribution
Poisson distribution
Binomial distribution

...



e.g. bear body size



Bayesian

generalised mixed models

with *MCMCglmm*

Bayesian *data analysis*

generalised

mixed models

with *MCMCglmm*

- Complex problems
- Inherent uncertainty that needs to be quantified
- Integration of different sources of information

BAYES THEOREM

Bayesian generalised mixed models with *MCMCglmm*

[Can you solve the false positive riddle?](#)
by Alex Gendler



TEDEd

Bayesian generalised mixed models with *MCMCglmm*

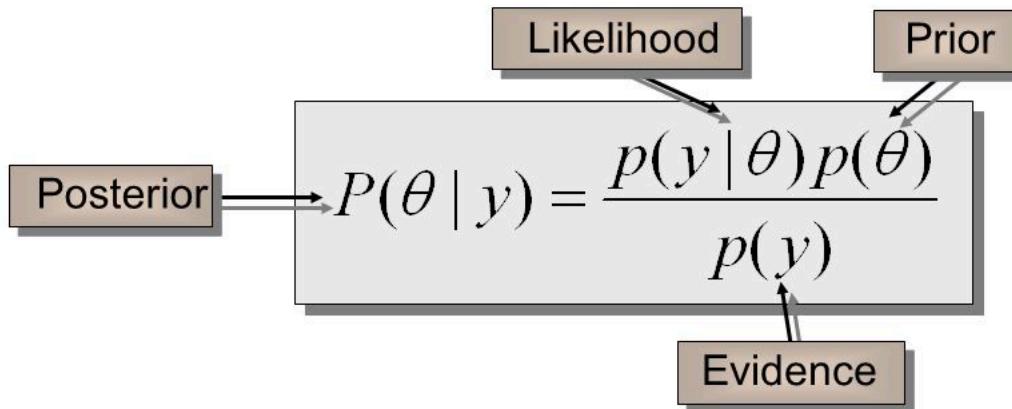
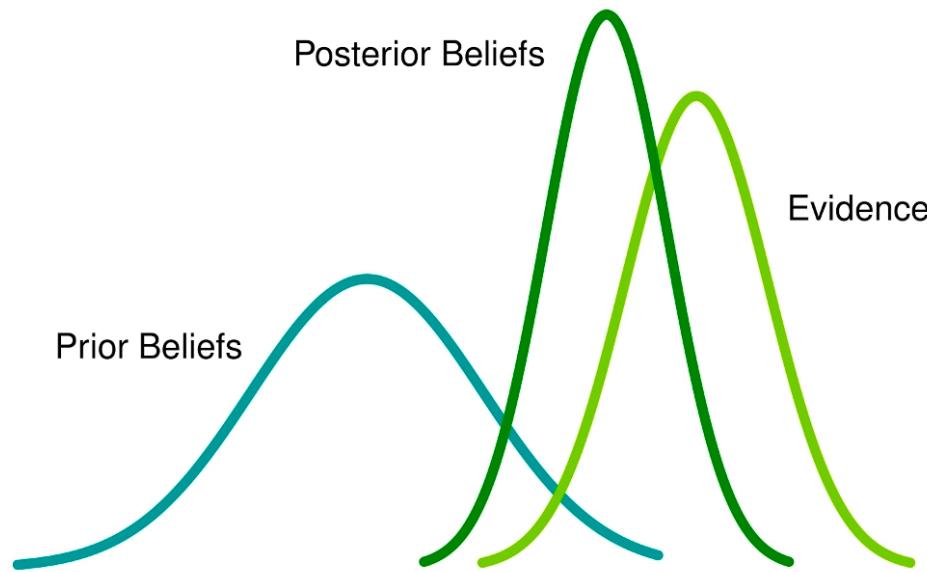


BAYES THEOREM



[Monty Hall Problem](#)
(extended math version)

Hint: Always switch!



The **posterior probability** is proportional to **prior probability** multiplied by the **likelihood**. The frequentist approach only deals with the likelihood.

Bayesian generalised mixed models with *MCMCglmm*

Advantages:

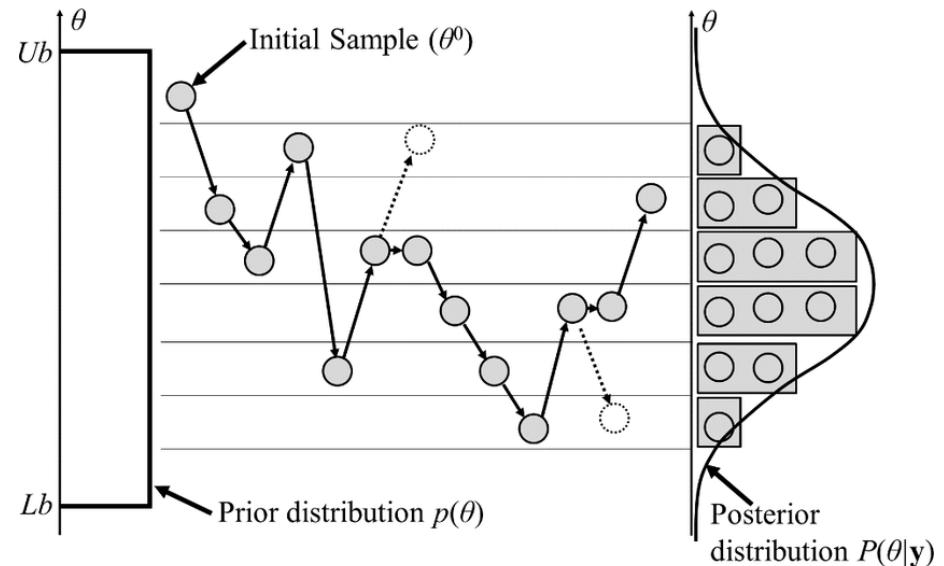
- conceptually coherent (we can test a hypothesis given the data)
- very complex models possible

Disadvantages:

- not always clear how to determine the priors
- computationally intensive (e.g. Markov chain Monte Carlo algorithm)

Bayesian generalised mixed models with *MCMCglmm*

Markov chain Monte Carlo



Bayesian generalised mixed models with *MCMCglmm*

Chocolate hills example

<https://www.youtube.com/watch?v=XV4yj4T4PBQ>



Tutorial: Pigeons & Doves

Example data from Lapiendra *et al.*, (2014) *Proc Royal Soc B*

<https://doi.org/10.1098/rspb.2012.2893>

Terrestrial species



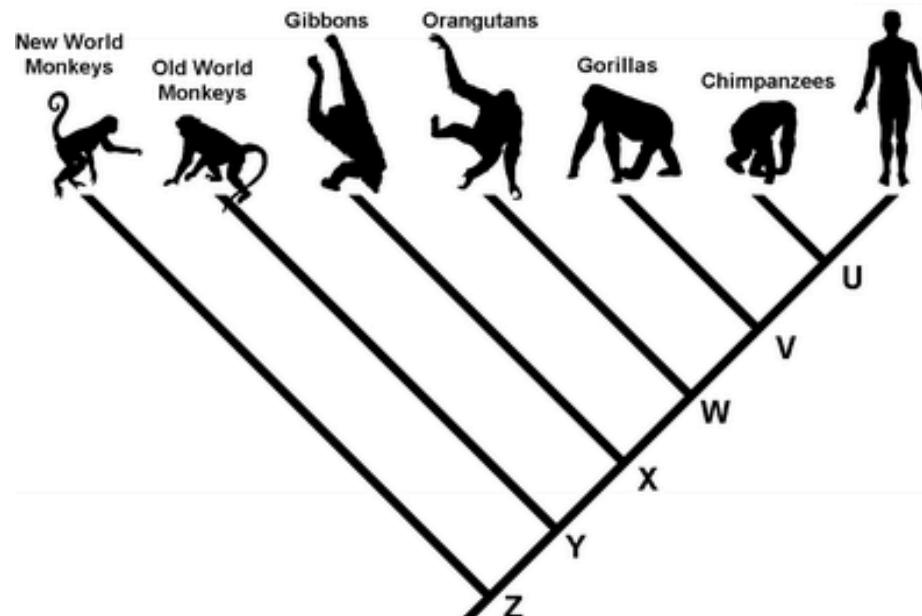
Arboreal species



Phylogenetic autocorrelation

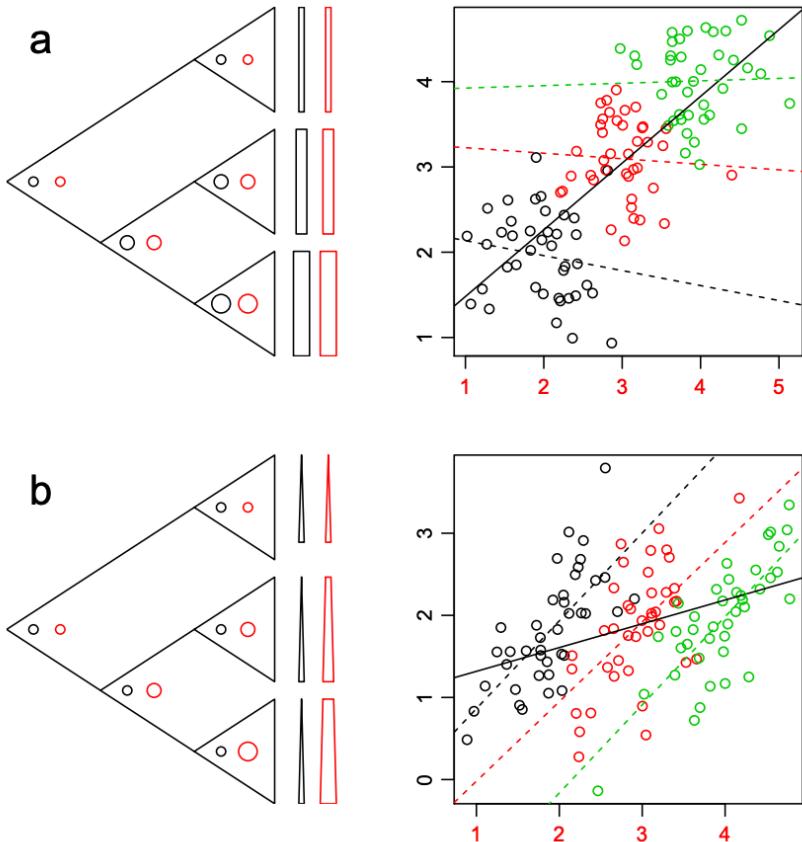
"Everything is related to everything else, but near things are more related than distant things".

Waldo R. Tobler



Phylogenetic autocorrelation

caper R-package (Orme, D. et al., (2013))



Simple regression (solid line) suggest strong relationship

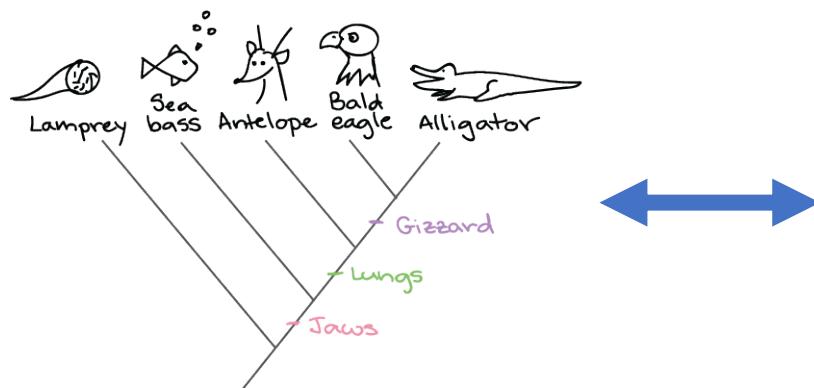
No consistent relationship (dashed lines) within groups.

Simple regression (solid line) suggests weak relationship

Strong positive relationships between the traits within groups.

Phylogenetic autocorrelation in MCMCglmm

What we need?



animal	jaws	lungs
Lamprey	no	no
Sea bass	yes	no
Antelope	yes	yes
Bald eagle	yes	yes
Alligator	yes	yes

- A phylogenetic tree (*pedigree*)
- Names in data matching tip labels (*animal*)

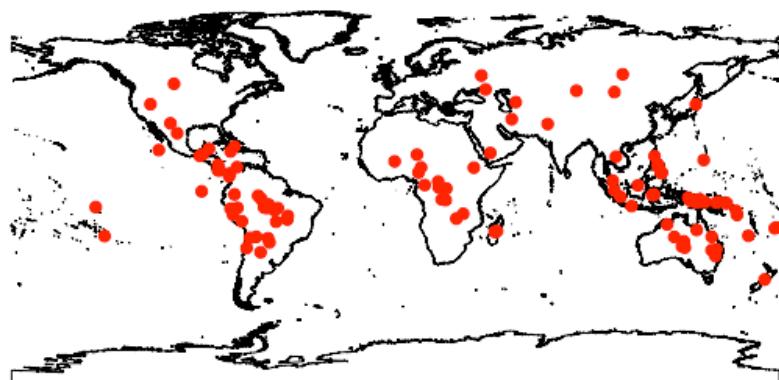
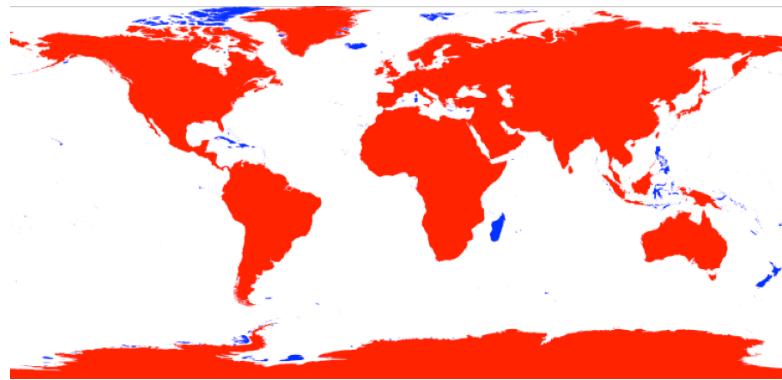
Combining Spatial analysis and MCMCglmm

- We can infer some species characteristics with spatial analysis
- We can then use this to run our analysis in MCMCglmm

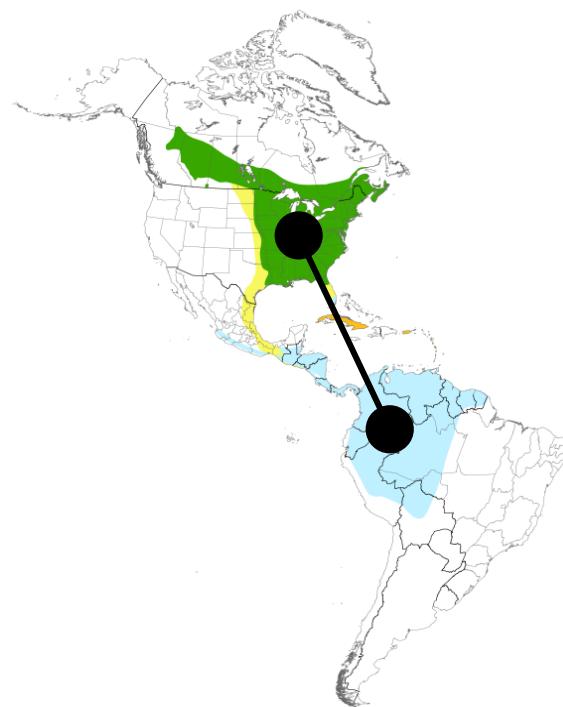


Combining Spatial analysis and MCMCglmm

Island species



Migratory species



Some additional information about priors

For fixed factors we will use a diffuse normal prior centred around zero but with very large variance (10^8).

For variance components, prior specification $V=1$ and $\text{nu}=0.002$ is frequently used.

where `priorR` is a list with elements `V` and `nu` specifying the prior for the variance, and `priorB` is a list with elements `mu` and `V` specifying the prior for the mean. `MCMCglmm` takes these prior specifications as a list:

```
> prior <- list(R = list(V = 1, nu = 0.002), B = list(mu = 0,
+     V = 1e+08))
```

Univariate inverse Wishart with the variance at the limit set to 1 ($V=1$) and varying degree of belief parameter (nu). With $V=1$ these distributions are equivalent to inverse gamma distributions with shape and scale parameters set to $\text{nu}/2$.

Hadfield, J. (2012). MCMCglmm course notes.

<http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>.

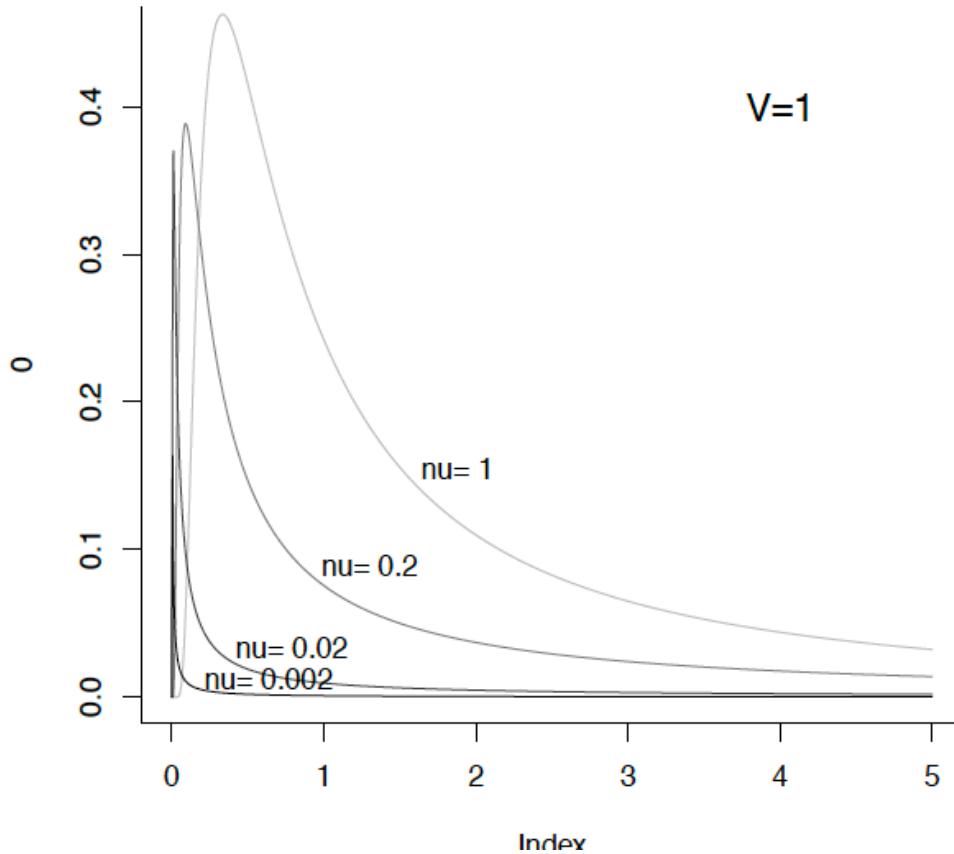


Figure 1.5: Probability density function for a univariate inverse Wishart with the variance at the limit set to 1 ($V=1$) and varying degree of belief parameter (ν). With $V=1$ these distributions are equivalent to inverse gamma distributions with shape and scale parameters set to $\nu/2$.

Hadfield, J. (2012). MCMCglmm course notes.

<http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>.