The Future of Bayes and brms

With Applications in Biology

Paul Bürkner

In Bayes we quantify uncertainty with probability

Rethinking the Bayes Theorem

$$p(\theta \mid y) = \frac{p(y \mid \theta)\, p(\theta)}{p(y)}$$
$$\propto p(y \mid \theta)\, p(\theta)$$
$$= p(y, \theta)$$

What I like and don't like about Bayesian inference

What I like:

- Intuitive approach to expressing uncertainty
- Ability to incorporate prior information
- A lot of modeling flexibility
- Joint posterior distribution of parameters
- Easy propagation of uncertainty

What I don't like:

- Slow Speed of model estimation

Phylogenies and pedigrees

Both phylogenies and pedigrees induce dependencies between individuals

Varying coefficients ("random effects") with an exchangeable prior:

$$\theta_i \sim \text{normal}(0, \sigma_\theta^2)$$

Varying coefficients with a non-exchangeable, joint prior:

$$\theta \sim \text{normal}(0, \sigma_\theta^2 \, A)$$

A is a covariance matrix implied by a phylogeny or pedigree

Phylogenies and pedigrees

Simple simulated dataset with a phylogenetic structure

| phen	cofactor	phylo
107.07	10.31	sp_1
79.61	9.69	sp_2
116.38	15.01	sp_3
143.29	19.09	sp_4
139.61	15.66	sp_5
68.51	6.01	sp_6

Reference: de Villemeruil P. & Nakagawa, S. (2014) General quantitative genetic methods for comparative biology. In: *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice.* Springer.

Phylogenies and pedigrees: Model

```
fit_phylo <- brm(
  phen ~ cofactor + (1 | gr(phylo, cov = A)),
  data = data_phylo, data2 = list(A = A),
  family = gaussian(),
  prior = prior(normal(0, 10), "b") +
    prior(normal(0, 50), "Intercept") +
    prior(student_t(3, 0, 20), "sd"),
  file = "models/fit_phylo"
)
```
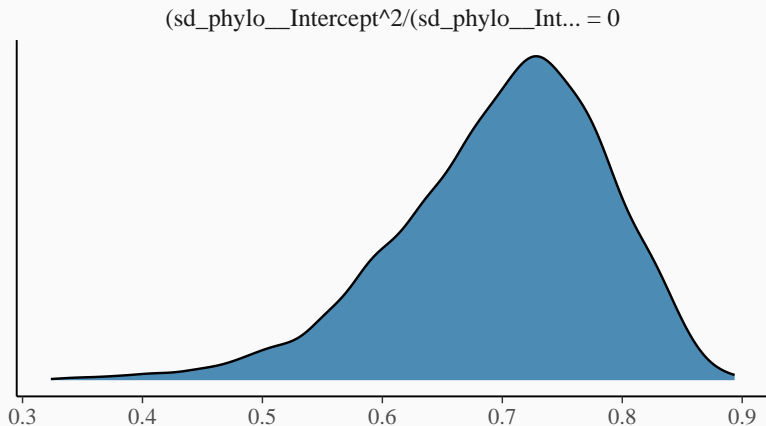
```
plot(conditional_effects(fit_phylo, "cofactor"), points = TRUE)
```
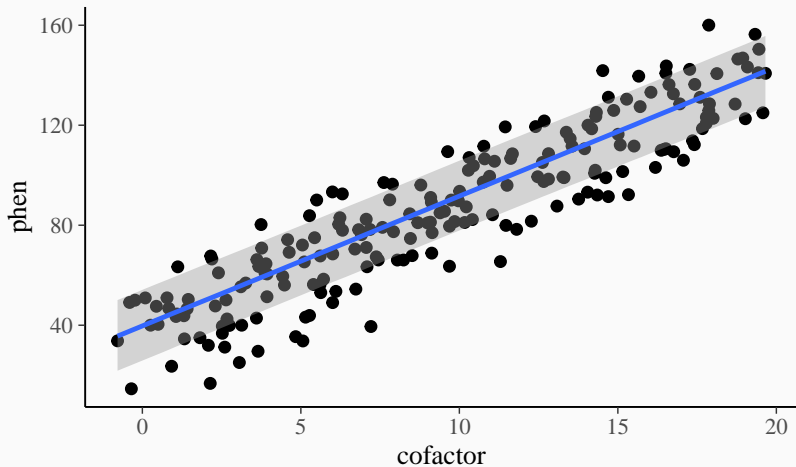
Phylogenies and pedigrees: Explained variance

```
hyp <- "sd_phylo__Intercept^2 / (sd_phylo__Intercept^2 + sigma^2) = 0"
hyp <- hypothesis(fit_phylo, hyp, class = NULL)
plot(hyp)
```


$(sd\_phylo\_\_Intercept^2/(sd\_phylo\_\_Int... = 0$

Multivariate model: Phenotype of blue tits

Study Objective: Predicting tarsus length and back color reflection of chicks based on environmental and genetic factors

Environmental factor: Fosternest

Genetic factor: Genetic mother (dam)

Control variables: Hatchdate and sex of the chicks

Reference: Hadfield, J. D., Nutall, A., Osorio, D., & Owens, I. P. F. (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. Journal of evolutionary biology, 20(2), 549-557.

Phenotype of blue tits: Data Overview

tarsus	back	dam	fosternest	hatchdate	sex
-1.89	1.15	R187557	F2102	-0.69	Fem
1.14	-0.76	R187559	F1902	-0.69	Male
0.98	0.14	R187568	A602	-0.43	Male
0.38	0.26	R187518	A1302	-1.47	Male
-0.08	-0.30	R187528	A2602	-1.47	Fem
-1.14	1.56	R187945	C2302	0.35	Fem
-1.14	-0.43	Fem3	C1902	-0.43	Male
1.89	-1.34	R187030	C1302	-0.95	Fem
-0.38	0.07	R187517	C602	-1.98	Fem
-0.08	-0.12	R187523	B2202	-0.95	Fem

Multivariate models in brms

```r
bf_tarsus <- bf(
  tarsus ~ sex + (1 | gr(fosternest, id = "p")) +
    (1 | gr(dam, id = "q")))

bf_back <- bf(
  back ~ hatchdate + (1 | gr(fosternest, id = "p")) +
    (1 | gr(dam, id = "q")))

fit_mv <- brm(
  bf_tarsus + bf_back + set_rescor(TRUE),
  data = BTdata,
  chains = 4, cores = 4,
  file = "models/fit_mv"
)
```

Group-level standard deviations and correlations:

dam	Estimate	l-95% CI	u-95% CI
sd(tarsus_Intercept)	0.48	0.39	0.59
sd(back_Intercept)	0.25	0.12	0.39
cor(tarsus_Intercept,back_Intercept)	-0.50	-0.90	-0.07

fosternest	Estimate	l-95% CI	u-95% CI
sd(tarsus_Intercept)	0.27	0.16	0.38
sd(back_Intercept)	0.34	0.23	0.46
cor(tarsus_Intercept,back_Intercept)	0.69	0.22	0.99

Gaussian process modeling

Exact Gaussian processes may be slow for bigger data:

```r
fit_gp <- brm(
  tarsus ~ gp(back) + (1 | fosternest) + (1 | dam),
  data = BTdata,
  chains = 4, cores = 4,
  file = "models/fit_gp"
)
```

After 15 minutes the model wasn't even at 10%
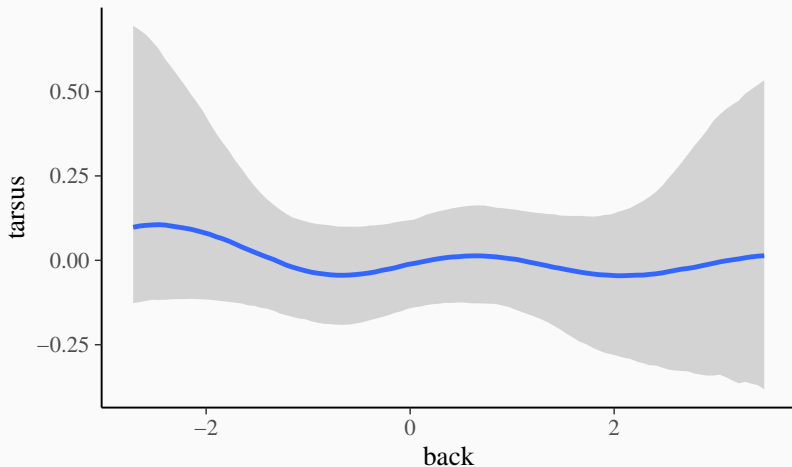
Gaussian process modeling

Approximate Gaussian processes will be much faster:

```
fit_hsgp <- brm(
  tarsus ~ gp(back, k = 10) +
    (1 | fosternest) + (1 | dam),
  data = BTdata,
  chains = 4, cores = 4,
  control = list(adapt_delta = 0.99),
  file = "models/fit_hsgp"
)
```

Only took 20 seconds per chain

Approximate Gaussian processes: Results

```
conditional_effects(fit_hsgp)
```

Modeling missing values

Current functionality in brms 2.22:

```
bf(food ~ mi(age) + covariates + (1 | ape)) +
  bf(age | mi() ~ covariates)
```

Extended functionality in brms 3.0:

```
bf(food ~ mi(age) + covariates + (1 | ape)) +
  bf(age | mi(idx = ape) ~ covariates,
      family = lognormal())
```

Current dev version available via:

```
remotes::install_github("paul-buerkner/brms", ref = "brms3")
```

Structural equation modeling

The effect of personality in primates on their hierarchical position:

Planned functionality in brms 3.0:

```
bf(hierachy ~ sex * lv(agreeableness)) +
bf(agreeableness | lv() ~ 1) +
  bf(indicator1 ~ lv(agreeableness),
     family = poisson()) +
  bf(indicator2 ~ lv(agreeableness),
     family = poisson())
```

## Overview of key changes in brms 3.0

Key features:

- New structural equation modeling
- Extend missing value modeling
- Extend multilevel modeling
- Parameter sharing in multivariate models
- Uncertainty propagation during post-processing
- Allow installation without rstan

Other changes:

- Refactor internal model representations
- Tighter integration of the posterior package
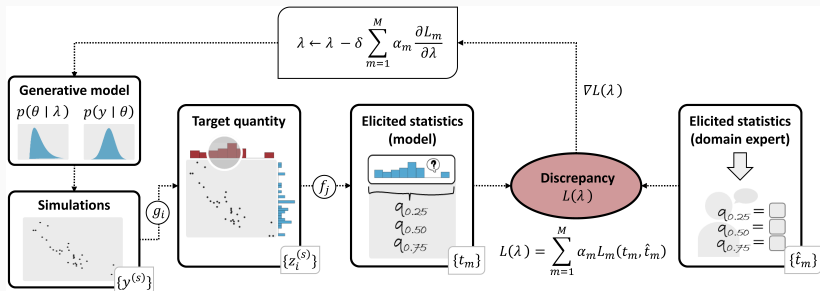- Remove many deprecated features and arguments

A brief look into my own research
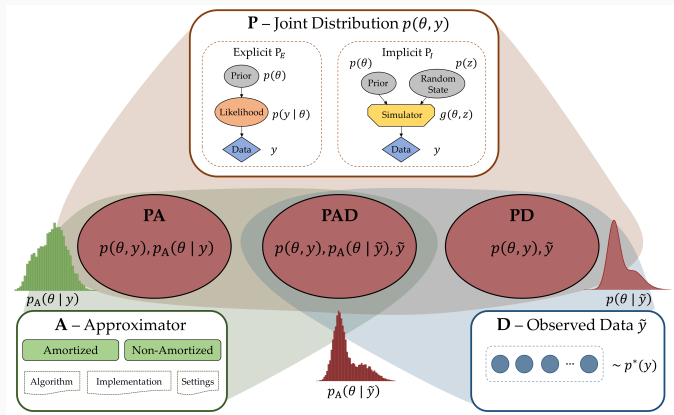
Further reading: Bürkner P. C., Scholz M., & Radev S. T. (2023). Some models are useful, but how do we know which ones? Towards a unified Bayesian model taxonomy. *Statistics Surveys*. doi:10.1214/23-SS145
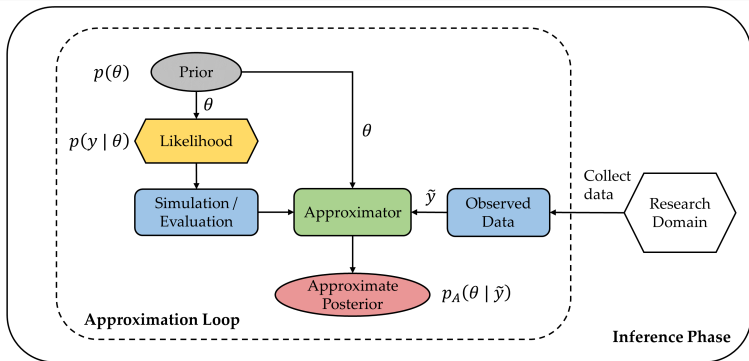
How can we incorporate expert knowledge into our models?

Further reading: Bockting F., Radev S. T., & Bürkner P. C. (2024). Simulation-Based Prior Knowledge Elicitation for Parametric Bayesian Models. *Scientific Reports*. doi:10.1038/s41598-024-68090-7
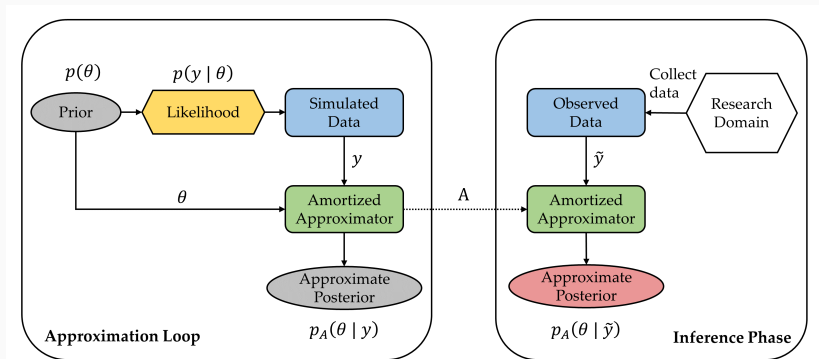
How can we improve the standard inference setting?

Further reading: Vehtari A., Gelman A., Simpson D., Carpenter B., & Bürkner P. C. (2021). Rank-normalization, folding, and localization: An improved Rhat for assessing convergence of MCMC (with discussion). *Bayesian Analysis*. doi:10.1214/20-BA1221

How far can we scale amortized inference?

Further reading: Radev S. T., Schmitt M., Schumacher L., Elsemüller L., Pratz V., Schälte Y., Köthe U., & Bürkner P. C. (2023). BayesFlow: Amortized Bayesian Workflows With Neural Networks. *Journal of Open Source Software*. doi:10.21105/joss.05702

The future of Bayes in the era of deep learning

Deep learning for Bayes vs. Bayes for deep learning

More about me and my research

- Website: https://paulbuerkner.com/
- Email: paul.buerkner@gmail.com
- Bluesky: @paulbuerkner.com