University of Zurich[UZH]

Institute of Psychology – Cognitive Psychology

Bayesian Modelling for Observational Data

Workshop for the R Peer Mentoring Group

Gidon Frischkorn (with special thanks to Philipp Musfeld)

Spring Term 2024

Today’s Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

--- Break ---

3. Bayesian Regression Models using Stan (brms)
4. Statistical Inference in Bayesian Models

--- Break ---

5. Practical Exercises

Today's Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
    - What is the difference between a Bayesian and a Frequentist?
    - Unpacking Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
3. Bayesian Regression Models using Stan (brms)
4. Statistical Inference in Bayesian Models
5. Practical Exercises

1. What is the difference between a Bayesian and a Frequentist?

Frequentist	Bayesian
A hypothesis is either true or false	Hypothesis are probably true or false
We can discover truth with sufficiently large samples & large number of experimentes	We update our beliefs (i.e. the probabilities) in hypothesis by collecting data
Quantifies the probability of observing data or more extreme data assuming the null hypothesis is true	Quantifies the probability of a hypothesis being true given the observed data
Computationally inexpensive and very quick	Computationally expensive and slow (sometimes really slow…)

1. What is the difference between a Bayesian and a Frequentist?

A short thought experiment

We have designed a test for evaluating if a person is a genius. The test detects geniuses with 99% accuracy and provides negative results for normal people with 99.5% accuracy.

We tested a random person at Zurich HB and the result indicated that the tested person is a genius. What do we conclude?

Frequentist

→ Assuming the person is no genius (H_0), the result has a probability of 0.5% (p = .005), therefore the person is a genius ☺

Bayesian

→ I do not know; I need more information…

1. What is the difference between a Bayesian and a Frequentist?

Short thought experiment

We have designed a test for evaluating if a person is a genius. The test detects geniuses with 99% accuracy and provides negative results for normal people with 99.5% accuracy.

We tested a random person at Zurich HB and the result indicated that the tested person is a genius. What do we conclude?

Bayesian

→ I do not know; I need more information…

What is the information we have?

1. The probability of our data/evidence given the different outcomes: The **likelihood**

What is the additional information a Bayesian needs?

1. The probability of our outcome/hypothesis (before having seen the data): The **prior**
2. The probability of our data/evidence (given all possible outcomes): The **marginal likelihood**

→ Using these info and Bayes Theorem, we can quantify our belief in an outcome/a hypothesis given the data/evidence we collected: The **posterior**

$$P(H|D) = \frac{P(H) * P(D|H)}{P(D)}$$

$$P(D) = P(H) * P(D|H) + P(\neg H) * P(D|\neg H)$$

1. What is the difference between a Bayesian and a Frequentist?

Short thought experiment

We have designed a test for evaluating if a person is a genius. The test detects geniuses with 99% accuracy and provides negative results for normal people with 99.5% accuracy.

We tested a random person at Zurich HB and the result indicated that the tested person is a genius. What do we conclude?

Some additional assumptions:

- In 100.000 people there are about 200 geniuses.

$$P(H|D) = \frac{P(H) * P(D|H)}{P(D)}$$

	Before Testing Prior	At test Likelihood	After test Posterior
Genius	$\frac{200}{100.000} = 0.002$	0.99	$\frac{0.002 \times 0.99}{0.002 \times 0.99 + 0.998 \times 0.005} = 0.284$
No Genius	$\frac{99.800}{100.000} = 0.998$	0.005	$\frac{0.998 \times 0.005}{0.002 \times 0.99 + 0.998 \times 0.005} = 0.716$

1. What is the difference between a Bayesian and a Frequentist?

What are your questions?

1. What is the difference between a Bayesian and a Frequentist?

A short exercise

A shop for magicians' supplies sells fair and trick coins. The trick coins can be biased towards showing tails or heads in 80% of the flips. A friend got you a coin from this shop but did not tell you if it was fair or a trick coin.

So now we must find out for yourselves. The following steps should help:

1. Specify prior beliefs what kind of coin you might have
2. Find a distribution to simulate a coin flip given a certain probability
3. Simulate data for 5 flips for a coin randomly sampled from all possible outcomes
4. Evaluate how likely your results are given that the coin is fair or a trick coin
5. Use these likelihoods to update your prior beliefs
6. Repeat steps 3. to 5. for another 5 flips and update your belief again.

Today’s Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
   - What is a Sampler and why do we need them?
   - How does a (simple) sampler work?
3. Bayesian Regression Models using Stan (brms)
4. Statistical Inference in Bayesian Models
5. Practical Exercises

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

What is a Sampler and why do we need them?

Prior

Posterior

Likelihood

$$P(H|D) = \frac{P(H) * P(D|H)}{P(D)}$$

Marginal Likelihood

Often, we don't know this.
But why?

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

What is a Sampler and why do we need them?

Marginal Likelihood:

- The Probability of generating the observed data for ALL possible values of the model parameters
- *In more technical terms:* The Likelihood function when integrating over the whole parameter space

$$P(\Theta|D) = \frac{P(\Theta) * P(D|\Theta)}{P(D)} = \frac{P(\Theta) * P(D|\Theta)}{\int P(D|\Theta) * P(\Theta)\, d\Theta}$$

- *In more practical terms*: A constant which normalizes the Posterior distribution to make it a proper probability distribution (the area under the distribution has to integrate to 1)

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

What is a Sampler and why do we need them?

Example 1: Discrete space of possible parameter values:

We have designed a test for evaluating if a person is a genius. The test detect geniuses with 99% accuracy and provides negative results for normal people with 99.5% accuracy.

We tested a random person at Zurich HB and the result indicated that the tested person is a genius. What do we conclude?

$$P(G|T=1) = \frac{P(Genius=1) * P(Test=1|Genius=1)}{P(D)}$$

$$P(G|T=1) = \frac{P(G=1) * P(T=1|G=1)}{P(T=1|G=1) * P(G=1) + P(T=1|G=0) * P(G=0)}$$

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
What is a Sampler and why do we need them?

Example 2: Continuous space of possible parameter values:
We got a coin from a magician supply store. We have no idea whether the coin is fair, whether is has a bias, nor how big the bias could be. Thus, the probability θ for throwing heads could be any value between 0 and 1.

$$P(\theta = 0.5|D) = \frac{P(\theta = 0.5) * P(D|\theta = 0.5)}{P(D)} = \frac{P(\theta = 0.5) * P(D|\theta = 0.5)}{\int P(D|\Theta) * P(\Theta)\, d\Theta}$$

→ *To know the marginal likelihood, we must integrate the likelihood function over the continuous parameter space*

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

What is a Sampler and why do we need them?

What does a Sampler do?

- The marginal likelihood is a normalization constant. It is required to make the posterior distribution a proper probability distribution
- Without knowing the marginal likelihood, we can still obtain the ***unnormalized posterior distribution:***

$$unP(\Theta|D) = P(\theta) * P(D|\Theta) \qquad \text{or} \qquad P(\Theta|D) \propto P(\theta) * P(D|\Theta)$$

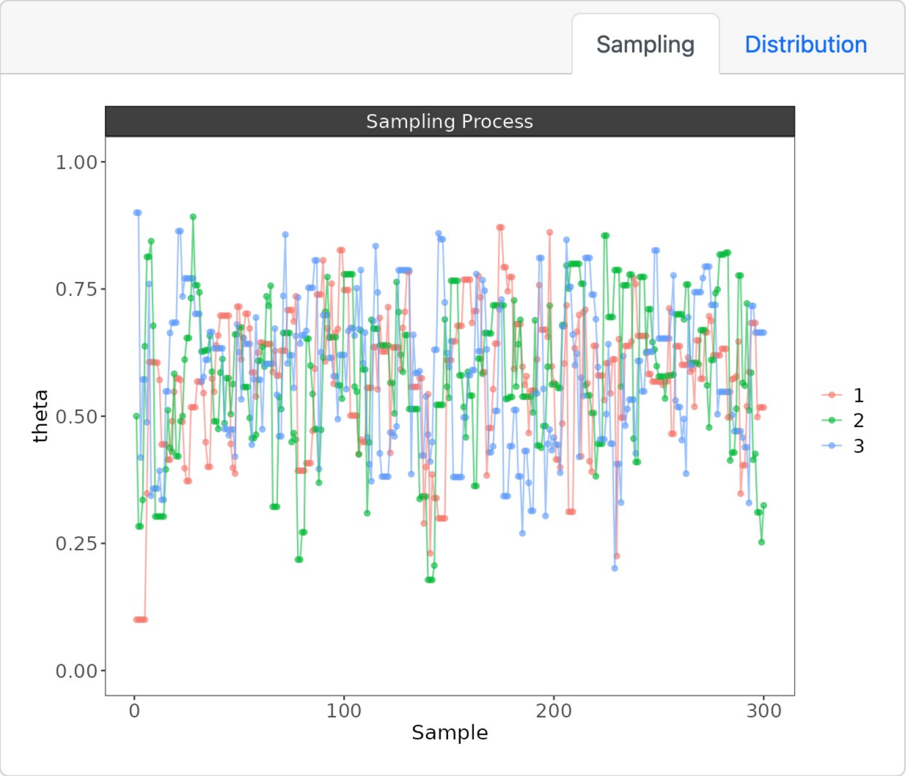
- The unnormalized posterior distribution has the exact same ***shape*** as the normalized posterior distribution (just on a different scale)
- By drawing a large number of samples from the unnormalized posterior distribution, we can approximate the normalized posterior distribution, without knowing the marginal likelihood

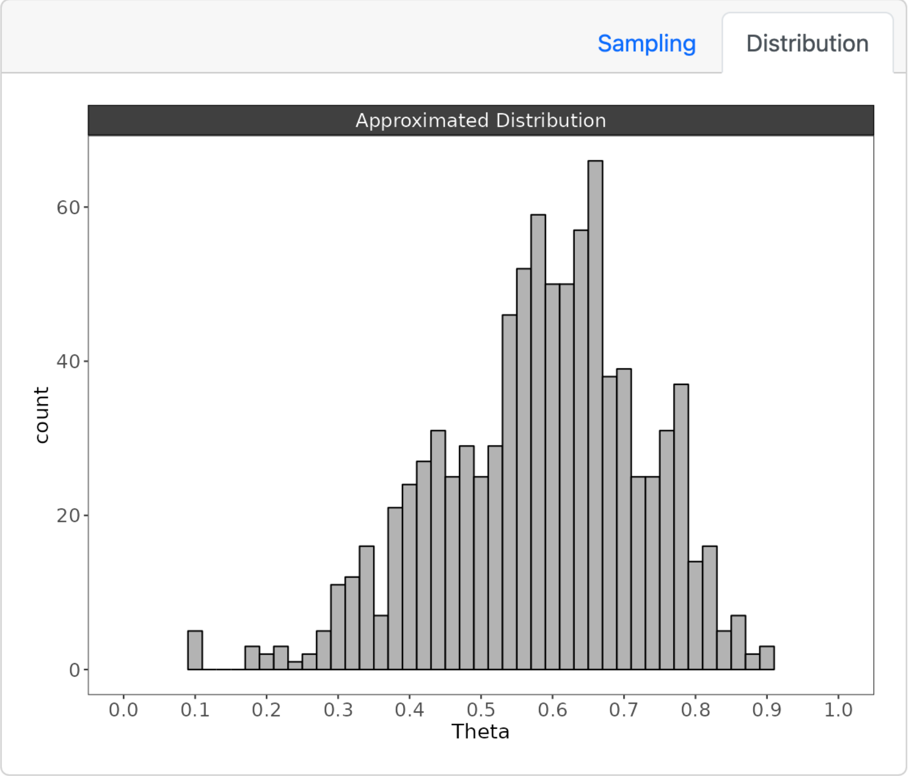
→ This is what a sampler does: It draws samples from the unnormalized posterior distribution to approximate the normalized posterior distribution, while ignoring the marginal likelihood

2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

How does a (simple) sampler work?

Let's have a look at an example: https://pmusfeld.shinyapps.io/Bayesian_Posterior_Sampling/





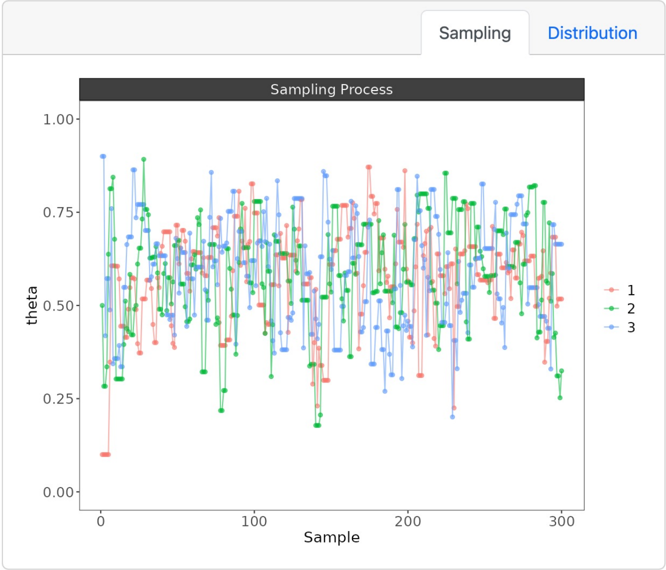
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

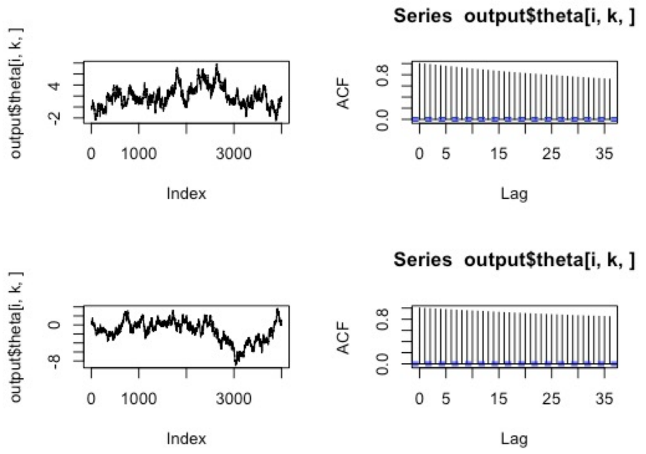
How does a (simple) sampler work?

Checking if sampling was successful

1. **Convergence**
   - did all chains (that started from different values) arrive at the same parameter estimate
   - R-hat = ratio of within chain variance / between chain variance → 1.00
2. **Auto-Correlation / Effective Sample size**
   - How strongly are samples correlated with the previous sampling steps → ideally this should be low
   - Effective Sample size = number of samples adjusted for auto-correlation between sampling steps
3. **Other convergence issues** (specific to the sampler in STAN
   1. Divergent transitions = unusually large jumps during sampling (bad!)
   2. reaching the maximum treedepth

→ Can be accommodated by adjusting sampler settings





2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?

What are your questions?

Today’s Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
3. Bayesian Regression Models using Stan (brms)
    - A short intro to probabilistic models
    - brms: An easy way to specify (almost) any model
4. Statistical Inference in Bayesian Models
5. Practical Exercises

3. Bayesian Regression Models using Stan (brms)

A short intro to probabilistic models

A short intro to probabilistic models

→ provides a formalized story of where the data can stem from

An example: We want to compare if the food portions handed out by Dora are smaller or larger than the portions handed out by Mario. To do so, we collect a random sample of portions by both Dora & Mario and compare them.

What do we need to know for specifying a formalized model for this example?

- The type of observation we make → weight, volume, count (a continuous positive variable)
- data generating process → some form of distribution
- Parameters → the variables determining the shape and location of the distribution

3. Bayesian Regression Models using Stan (brms)

A short intro to probabilistic models

Some background info on notation

- Observed/random variables are denoted with Latin letters: X, Y, etc.
- Parameters are denoted with Greek letters: μ, σ, λ
- Distributions are denoted with short acronyms: $\mathcal{N}$ (Normal distribution) or $\mathcal{B}$ (Bernoulli, Binomial distribution)
- The ~ denotes that something «stems» or «is drawn» from a distribution, for example:

$$X \sim \mathcal{N}(\mu, \sigma)$$

means, the random variable X follows a normal distribution with a mean μ and a standard deviation σ.

3. Bayesian Regression Models using Stan (brms)

A short intro to probabilistic models

An example: We want to compare if the food portions (P) handed out by Dora are different from the portions handed out by Mario. To do so, we collect a random sample of portions by both Dora & Mario and compare them.

$$P_M \sim \mathcal{N}(\mu_M, \sigma_M)$$
$$P_D \sim \mathcal{N}(\mu_D, \sigma_D)$$

Parameters: μ_M, μ_D, σ_M, σ_D

Question: μ_M = μ_D ?

3. Bayesian Regression Models using Stan (brms)

A short intro to probabilistic models

Rerformulating the model in terms of a regression?

$$P_i \sim \mathcal{N}(\mu_0 + \beta_1 \times Y_i, \sigma)$$

The Portion *P* you get from the person *i* stems from a Normal distribution,

where the mean is a composite of some **average portion μ_0** that gets adjusted by the **difference β_1** for **different persons Y_i** that serve you,

And a **variability of portions σ** (residual variance) that is assumed to be constant across people that serve.

Nice sideeffect: We can use such models to simulate data for our experiment assuming a set of parameters ☺ we just need to use the random generation function of the distributions we used in our statistical model.

→ `rnorm(n, mean, sd)`

- n = number of random samples to generate
- Mean = mean of the normal
- SD = standard deviation of the normal

3. Bayesian Regression Models using Stan (brms)

What are your questions?

3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms (**B**ayesian **r**egression **m**odels using **S**tan)

What does the brms R package provide?

I. interface to Stan → fit Bayesian generalized linear models

II. formula syntax similar to lme4 → provides familiar & simple regression analyses.

III. wide range of response distributions supported → fit wide range of data

IV. lots of further modeling options:
- non-linear and smooth terms,
- auto-correlation structures,
- censored data,
- missing value imputation,
- and quite a few more …

V. all parameters of response distribution can be predicted (means, standard deviations, etc.)

VI. flexible prior specifications → encourages users to apply prior distributions that reflect their beliefs.

VII. Model fit can easily be assessed → posterior predictive checks, cross-validation, and Bayes factors.

3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

`brm` → main function to fit models using brms

required arguments:

1. formula → specifies the regression model to estimate
2. data → data set that contains all variables (**important:** match variable names to names in formula!)

defaults:

3. family = gaussian() → which probability distribution does the DV stem from
4. priors
5. sampler settings (number of warmup & post-warmup sample, number of MCMC chains, etc.)

```
brm(formula = DV ~ 1 + IV + (1 + IV | ID),
    data = myData,
    # optional arguments
    family = gaussian(),
    prior = myPriors,
    iter = 4000, warmup = 1500,
    nChains = 4)
```

3. Bayesian Regression Models using Stan (brms)
A (short) intro to brms

Introduction to the `brmsformula` syntax

```
response | aterms ~ pterms + (gterms | group)
```

1. the dependent variable
2. additional response information terms (truncation, number of trials, decision information)
3. predictor terms
4. group specific terms (random effects)
5. grouping variables

Important: By default, this formula predicts the primary distribution parameter → mean, location, or probability parameters

You can specify additional formulae to predict other distributional parameters (e.g. sigma), like:

```
dist_parm ~ pterms + (gterms | group)
```

3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

Introduction to the `brmsformula` syntax

`response | aterms ~ pterms + (gterms | group)`

1. As `pterms` and `gterms` you can use:
   - Categorical predictors, e.g. factors (mind their contrast coding!)
   - Continous predictors
2. As `group` variables you can use:
   - Integer IDs,
   - Factors
   - additionally, you can distinguish different groups (e.g., younger vs. Older adults) using an additional grouping variable
3. By default the formula includes an `Intercept` that you can code explicitly as `1` or suppress using `0`
4. You can include interactions between predictors:
   - Either using `*` (`pred1*pred2`) → includes both, main effects of the predictors & their interaction
   - Or using `:` (`pred1:pred2`) → includes only their interaction without their main effects

3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

Introduction to the `brmsformula` syntax

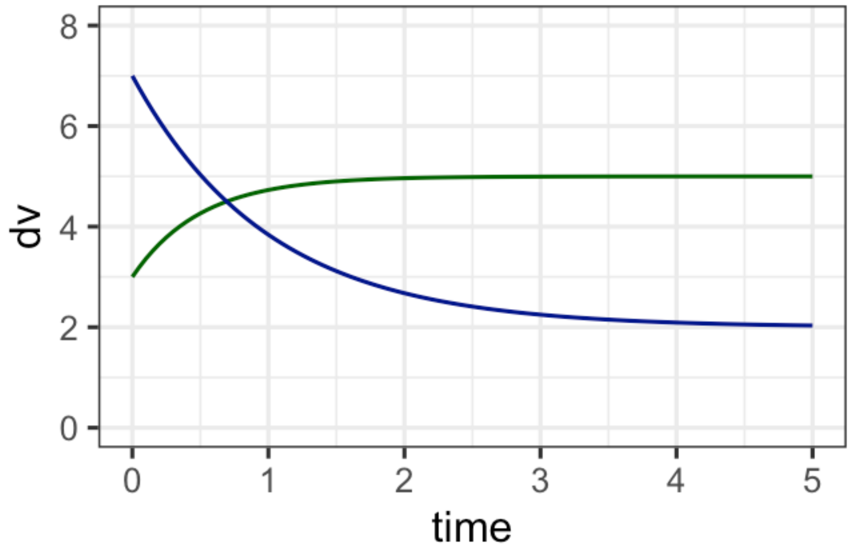
```
response | aterms ~ pterms + (gterms | group)
```

1. You can specify linear & non-linear formulae of arbitrary complexity. For example:

```
y ~ asymptote + (start - asymptote) * exp(-slope * time)
start ~ 1 + pred1 + (1 | ID)
asymptote ~ 1 + pred2 + (1 | ID)
slope ~ 1 + (1 | ID)
```

For this you, need to set the option for non-linear formulae to TRUE when setting up the formula, that is pass: `nl = TRUE`

Important: You should also consider setting adequate priors to ensure identification and proper convergence



for more info see the documentation.

3. Bayesian Regression Models using Stan (brms)

What are your questions?

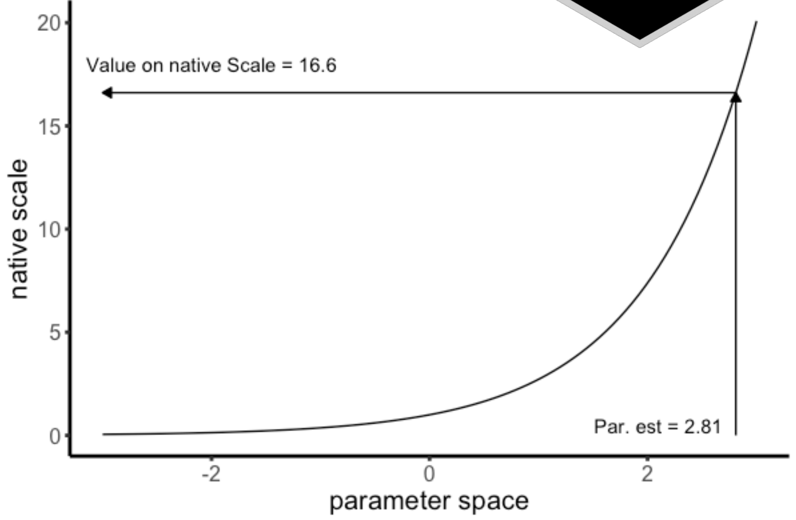
3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

Link Functions

Important: take care of link function and transformations when interpreting model parameters

→ for computational efficiency and ideal sampling `brms / Stan` transforms parameters with bounded parameter spaces (e.g., standard deviations > 0, 0 > probabilities < 1)

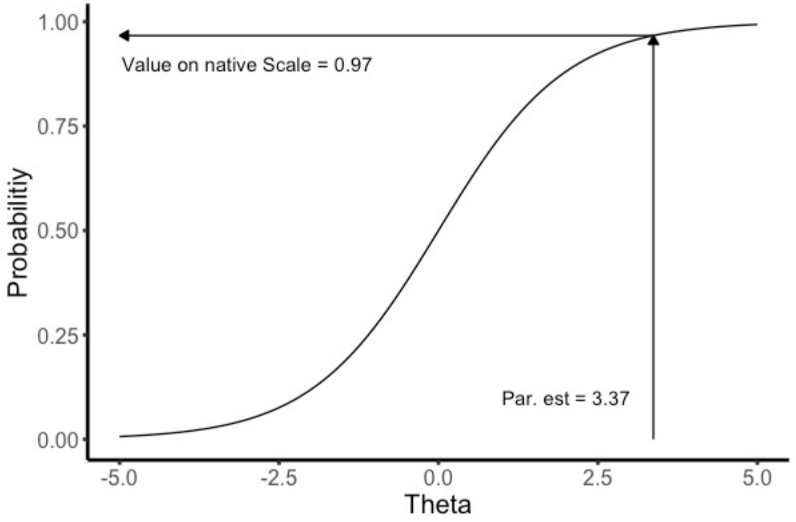


sd → log link

$$\sigma_{native} = e^{\sigma_{par}};\ \sigma_{par} = \log(\sigma_{native})$$

probabilities → logit link

$$p_{native} = \frac{e^{p_{par}}}{1 + e^{p_{par}}}$$



3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

Default priors

Can be accessed using the `default_prior` function, given a `formula` and `data`:

```
default_prior(my_formula, data, family = gaussian())
```

→ `brms` does not impose strong priors:
- no problem for parameter estimation,
- Bayes Factors & model comparison require user priors!

```
> default_prior(formula, ToothGrowth)
                  prior     class       coef group resp dpar nlpar lb ub       source
                 (flat)         b                                             default
                 (flat)         b      dose1                             (vectorized)
                 (flat)         b      dose2                             (vectorized)
                 (flat)         b      supp1                             (vectorized)
                 (flat)         b supp1:dose1                            (vectorized)
                 (flat)         b supp1:dose2                            (vectorized)
 student_t(3, 19.2, 9) Intercept                                              default
    student_t(3, 0, 9)     sigma                                   0          default
```

User priors

→ You can set priors using the prior function:

```
prior(normal(0,1), class = b)
```

You need to pass at least:
- the prior distribution
- the parameter class it should be applied to

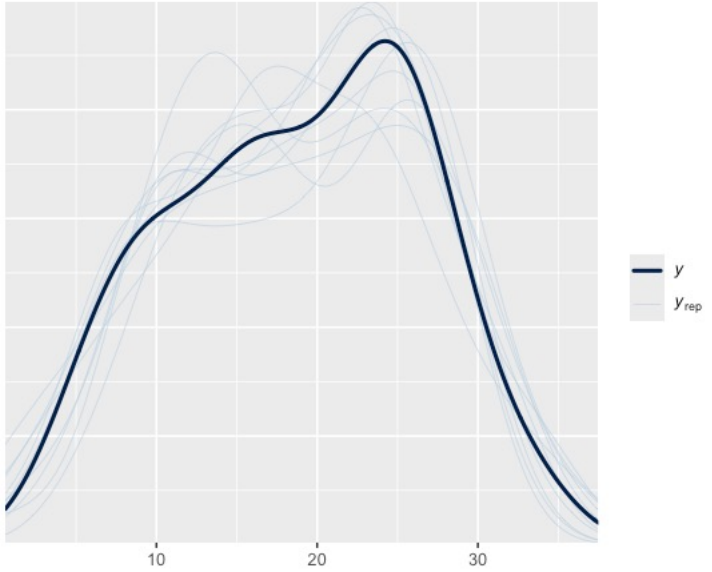
for more info see the documentation.

3. Bayesian Regression Models using Stan (brms)

A (short) intro to brms

Accessing and Evaluating model results

1. `pp_check(brmsfit)` → posterior predictive plot to visually evaluate model fit
2. `summary(brmsfit)` → overview of the estimated parameters
3. `fixef(brmsfit)` & `ranef(brmsfit)` → get fixed and random effect estimates
4. `conditional_effects(brmsfit)` → obtain and plot conditional effects for estimated brms models

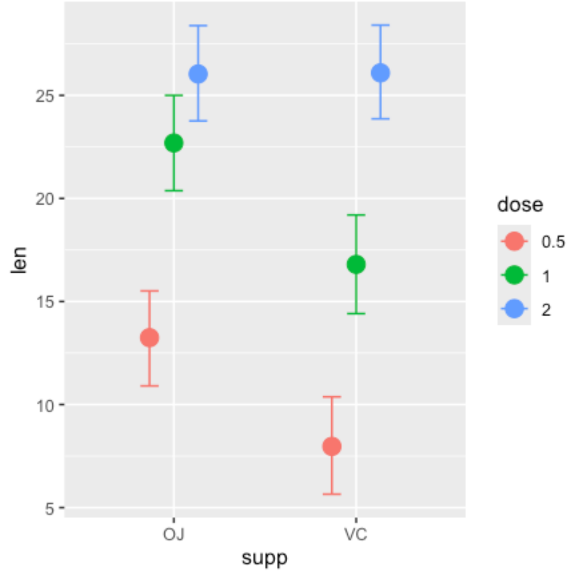



```
> summary(brmsfit)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: len ~ 1 + supp + dose + supp:dose
   Data: ToothGrowth (Number of observations: 60)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Regression Coefficients:
            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept      18.81      0.48    17.88    19.74 1.00     5921     2755
supp1          -2.62      0.67    -3.91    -1.27 1.00     6032     2592
dose1           4.50      0.82     2.92     6.12 1.00     5413     2984
dose2         -10.05      0.81   -11.66    -8.45 1.00     6676     2922
supp1:dose1     2.99      1.17     0.65     5.19 1.00     5849     3071
supp1:dose2    -1.35      1.17    -3.66     0.90 1.00     6928     2751

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     3.70      0.36     3.07     4.45 1.00     5218     3222

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```



3. Bayesian Regression Models using Stan (brms)

What are your questions?

Today's Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
3. Bayesian Regression Models using Stan (brms)
4. Statistical Inference in Bayesian Models
   - Evaluating Parameter estimates & Credibility Intervals
   - Bayes Factors & Model Comparison
5. Practical Exercises

4. Statistical Inference in Bayesian Models

Credible intervals (CI)

- describe and summarise **the uncertainty** in parameter estimates
- given the observed data, the effect has *CI width*% probability of falling within this range

If the CI does not contain a value (e.g. 0), we can conclude that given the data this value is not probable given the chosen width of the CI
→ typically, we choose 95%

```
> summary(brmsfit)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: len ~ 1 + supp + dose + supp:dose
   Data: ToothGrowth (Number of observations: 60)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Regression Coefficients:
            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept      18.81      0.48    17.88    19.74 1.00     5921     2755
supp1          -2.62      0.67    -3.91    -1.27 1.00     6032     2592
dose1           4.50      0.82     2.92     6.12 1.00     5413     2984
dose2         -10.05      0.81   -11.66    -8.45 1.00     6676     2922
supp1:dose1     2.99      1.17     0.65     5.19 1.00     5849     3071
supp1:dose2    -1.35      1.17    -3.66     0.90 1.00     6928     2751

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     3.70      0.36     3.07     4.45 1.00     5218     3222

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

4. Statistical Inference in Bayesian Models

Bayes Factors & Model Comparions

→ Bayes Factors (BFs) are indices of *relative* evidence of one “model” over another.

Conceptually, the **Bayes Factor** tells us how the data we have collected will update our beliefs relative to the prior odds of two models.

$$\underbrace{\frac{P(M_1|D)}{P(M_2|D)}}_{\text{Posterior Odds}} = \underbrace{\boxed{\frac{P(D|M_1)}{P(D|M_2)}}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{P(M_1)}{P(M_2)}}_{\text{Prior Odds}}$$

4. Statistical Inference in Bayesian Models

Short thought experiment

We have designed a test for evaluating if a person is a genius. The test detects geniuses with 99% accuracy and provides negative results for normal people with 99.5% accuracy.

We tested a random person at Zurich HB and the result indicated that the tested person is a genius. What do we conclude?

Remember our example from the beginning

Was our test informative? Let's calculate the Bayes Factor:

$$\underbrace{\frac{P(M_1|D)}{P(M_2|D)}}_{\text{Posterior Odds}} = \underbrace{\frac{P(D|M_1)}{P(D|M_2)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{P(M_1)}{P(M_2)}}_{\text{Prior Odds}}$$

$$\frac{P(G)}{P(\neg G)} \times BF = \frac{P(G \mid T = 1)}{P(\neg G \mid T = 1)}$$

$$\frac{.002}{.998} \times BF = \frac{.284}{.716}; BF \approx 198 = \frac{.99}{.005}$$

	Before Testing Prior	At test Likelihood	After test Posterior
Genius	0.002	0.99	0.284
No Genius	0.998	0.005	0.716

As a rule of thumb, we consider:

- BF > 3 as moderate
- BF > 10 as strong evidence

in favour of one model over the other.

4. Statistical Inference in Bayesian Models

$$\underbrace{\frac{P(M_1|D)}{P(M_2|D)}}_{\text{Posterior Odds}} = \underbrace{\boxed{\frac{P(D|M_1)}{P(D|M_2)}}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{P(M_1)}{P(M_2)}}_{\text{Prior Odds}}$$

Bayes Factors & Model Comparisons

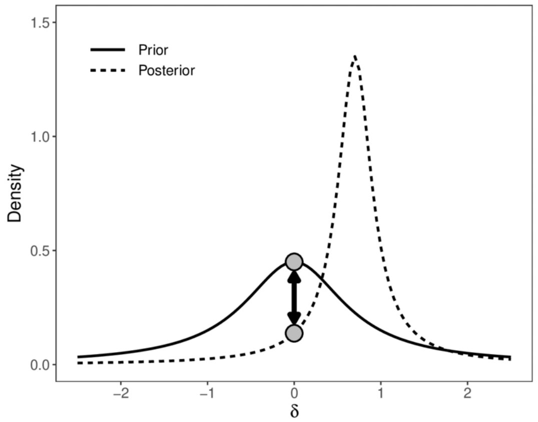
→ To indicate for which model the BF shows evidence for we use lower indices:

- BF_{12} = evidence in favor of Model 1 over Model 2
- From this we can calculate $BF_{21} = 1/BF_{12}$ → we can obtain evidence in favor of H_0

How can we estimate Bayes Factors?

1. brms includes a function `bayes_factor(brmsfit1, brmsfit2)` that calculates the BF_{12}
   - You need to set `save_pars = save_pars(all = T)` when fitting both models
   - works even for non-nested models
   - builds on `bridgesampling`, and thus you need many posterior samples (> 100 000)
2. alternatively, we can calculate BF for specific hypothesis (e.g. a parameter is equal to zero) using the `hypothesis` function:
   - You need to pass user priors to the model to obtain adequate Bayes Factors
   - you have to set the option `sample_prior = TRUE` when fitting the model

Savage-Dickey Density Ratio



for more info see the documentation: here and here

4. Statistical Inference in Bayesian Models

What are your questions?

Today’s Agenda

1. Introduction: Logic of Bayesian Statistics & Bayes Rule
2. Marcov-Chain Monte Carlo (MCMC) sampling – Why?
3. Bayesian Regression Models using Stan (brms)
4. Statistical Inference in Bayesian Models
5. Practical Exercises
    - Implement some analyses (own or example data)
    - Try obtaining evidence for or against hypothesis

5. Practical Exercises

Implement some analyses (own or example data)

Cardio Data

This dataset contains information on over 60.000 subjects for their age, height, weight as well as their blood pressure, smoking habits, and cardio desease status.

1. Get an idea how different variables (age, height, weight, smoking habits) are related to systolic (blood_high) blood pressure.
2. Find out which variables are risk factors for being diagnosed with a cadrio desease.

Learning experiment

This dataset contains information on a working memory learning experiment. Participants saw a lits of 9 letters they should remeber and this list was then repeated in 20 trials.

1. Implement a model that captures the learning effect over repetitions
2. Evaluate how including random effects changes the estimated effects

Bayesian Modelling for Observational Data

Thank you very much for your attention!

And welcome to the rabbit hole of Bayesian GLMs ☺

References

Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B. (2018). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic Bulletin & Review*, *25*(1), 219–234. https://doi.org/10/gc9mq5

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian Inference for Psychology. *Psychonomic Bulletin & Review*, *25*(1), 5–34. https://doi.org/10/gc9pqc

Oberauer, K. (2022). The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing. *Psychological Science*, *33*(4), 648–665. https://doi.org/10.1177/09567976211046884

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10/f4bx8f

Tendeiro, J. N., Kiers, H. A. L., Hoekstra, R., Wong, T. K., & Morey, R. D. (2024). Diagnosing the Misuse of the Bayes Factor in Applied Research. *Advances in Methods and Practices in Psychological Science*, *7*(1), 25152459231213371. https://doi.org/10.1177/25152459231213371

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, *60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10/gdz7tt

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10/gfgt6p