Introduction to Bayesian linear regression with brms — Part II: Bayesian Inference

Stefano Coretta

18/01/2020

Road map

1. Statistical inference.

Road map

1. Statistical inference.
2. Frequentist inference.

Road map

1. Statistical inference.
2. Frequentist inference.
3. Bayesian inference:

Road map

1. Statistical inference.
2. Frequentist inference.
3. Bayesian inference:
    3.1 Posterior distributions.

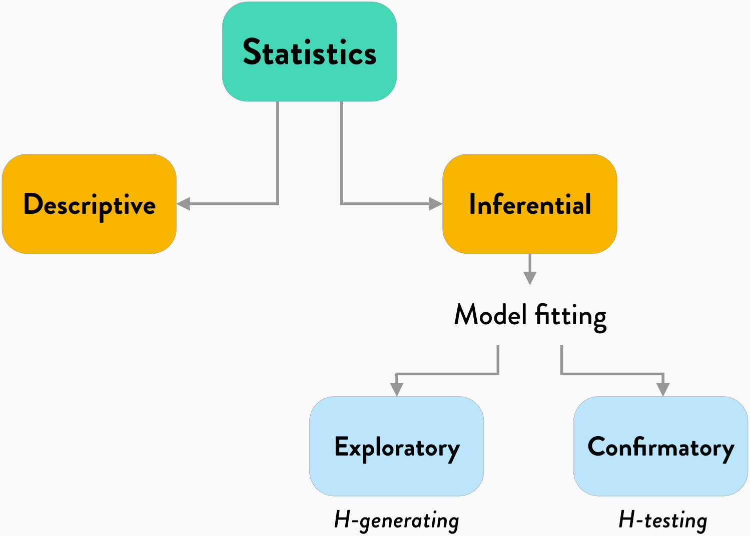
Road map

1. Statistical inference.
2. Frequentist inference.
3. Bayesian inference:
    3.1 Posterior distributions.
    3.2 ROPE.

Road map

1. Statistical inference.
2. Frequentist inference.
3. Bayesian inference:
    3.1 Posterior distributions.
    3.2 ROPE.
    3.3 Bayes factors.

STATISTICAL INFERENCE


Statistics
Descriptive
Inferential
Model fitting
Exploratory
Confirmatory
H-generating
H-testing

Hypothesis testing

We want to know two (interrelated) things:

- Is there **evidence** for the hypothesis H?

Hypothesis testing

We want to know two (interrelated) things:

- Is there **evidence** for the hypothesis H?
- What is the **strength** of the evidence?

Hypothesis testing

- We test H against empirical data (**hypothesis testing**).

Hypothesis testing

- We test H against empirical data (hypothesis testing).
    - It is important to decide in advance the details of the analysis.

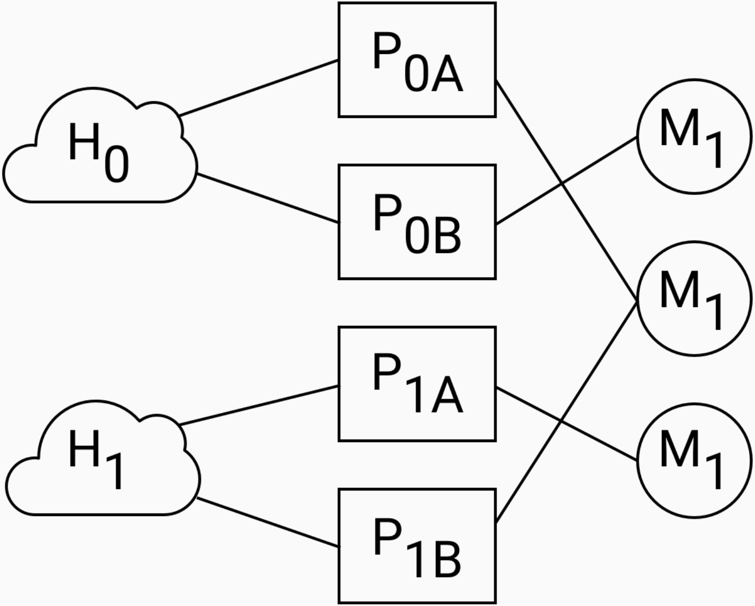
Hypothesis testing

- We test H against empirical data (**hypothesis testing**).
    - It is important to decide in advance the details of the analysis.
    - Even when you think you are not making decisions, the model is.

Hypothesis testing

- We test H against empirical data (**hypothesis testing**).
    - It is important to decide in advance the details of the analysis.
    - Even when you think you are not making decisions, the model is.
- Inference is ultimately a **long-term endeavour** (via accumulation of knowledge).

Hypothesis testing



FREQUENTIST INFERENCE

Frequentist statistics

- **Frequentist statistics** is based on the frequentist interpretation of probability.

Frequentist statistics

- **Frequentist statistics** is based on the frequentist interpretation of probability.
- **Frequentist probability of an event** is the *relative frequency* of occurrence of an event within an infinite set of imagined repetitions of that event.

Frequentist statistics

- **Frequentist statistics** is based on the frequentist interpretation of probability.
- **Frequentist probability of an event** is the *relative frequency* of occurrence of an event within an infinite set of imagined repetitions of that event.
- Most of modern science is based on *frequentism*.

Frequentist statistics

- **Frequentist statistics** is based on the frequentist interpretation of probability.
- **Frequentist probability of an event** is the *relative frequency* of occurrence of an event within an infinite set of imagined repetitions of that event.
- Most of modern science is based on *frequentism*.
    - lme4 package.

Frequentist statistics

- **Frequentist statistics** is based on the frequentist interpretation of probability.
- **Frequentist probability of an event** is the *relative frequency* of occurrence of an event within an infinite set of imagined repetitions of that event.
- Most of modern science is based on *frequentism*.
    - lme4 package.
    - **Null Hypothesis Significance Testing**.

Null Hypothesis Significance Testing

- NHST is based on falsification.

Null Hypothesis Significance Testing

- NHST is based on falsification.
    - Rejection of the Null Hypothesis (H_0).

Null Hypothesis Significance Testing

- NHST is based on falsification.
    - Rejection of the Null Hypothesis (H_0).
    - No direct “support/evidence” for hypotheses.

Null Hypothesis Significance Testing

- NHST is based on falsification.
    - Rejection of the Null Hypothesis (H_0).
    - No direct “support/evidence” for hypotheses.
- *P*-value (between 0 and 1).

Null Hypothesis Significance Testing

- NHST is based on falsification.
    - Rejection of the Null Hypothesis (H_0).
    - No direct “support/evidence” for hypotheses.
- *P*-value (between 0 and 1).
    - Probability of obtaining an estimate as extreme or more extreme, *assuming* H_0 *is true*.

Null Hypothesis Significance Testing

- NHST is based on falsification.
    - Rejection of the Null Hypothesis (H_0).
    - No direct “support/evidence” for hypotheses.
- *P*-value (between 0 and 1).
    - Probability of obtaining an estimate as extreme or more extreme, *assuming* H_0 *is true*.
    - Should be as low as possible.

Null Hypothesis Significance Testing

- Significance is dichotomous.

Null Hypothesis Significance Testing

- Significance is dichotomous.
  - $p < \alpha$ = “significant”.

Null Hypothesis Significance Testing

- Significance is dichotomous.
    - $p < \alpha$ = “significant”.
    - $p \geq \alpha$ = “non-significant”.

Null Hypothesis Significance Testing

- Significance is dichotomous.
  - $p < \alpha$ = "significant".
  - $p \geq \alpha$ = "non-significant".
- No degrees of significance.

Null Hypothesis Significance Testing

- Significance is dichotomous.
    - $p < \alpha$ = “significant”.
    - $p \geq \alpha$ = “non-significant”.
- No degrees of significance.
- “Significance” is a concept that makes sense only within frequentist statistics (NHST).

(Frequentist) confidence intervals

https://rpsychologist.com/d3/ci/

BAYESIAN INFERENCE

Bayesian inference

Bayesian statistics is based on the Bayesian interpretation of the **Bayes theorem**.

$$P(\theta \mid d) = \frac{P(d \mid \theta) \times P(\theta)}{P(d)}$$

Bayesian inference

- Evidence for any H (even the null).

Bayesian inference

- Evidence for any H (even the null).
- Strength of evidence.

Bayesian inference

- Evidence for any H (even the null).
- Strength of evidence.
- Capitalise on previous knowledge.

Bayesian inference

Three ways of doing hypothesis testing with Bayesian statistics:

- Inference from the **posterior**.

Bayesian inference

Three ways of doing hypothesis testing with Bayesian statistics:

- Inference from the posterior.
- Inference using a Region Of Practical Equivalence (ROPE).

Bayesian inference

Three ways of doing hypothesis testing with Bayesian statistics:

- Inference from the **posterior**.
- Inference using a **Region Of Practical Equivalence** (ROPE).
- Inference using the **Bayes factor**.

Inference from the posterior

1. Formulate a **hypothesis**:

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:
    - The posterior suggests that Condition B decreases reaction times by 15 to 80 ms at 95% confidence.

Inference from the posterior

1. Formulate a **hypothesis**:
   - H: Condition B decreases reaction times relative to Condition A.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
   - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:
   - The posterior suggests that Condition B decreases reaction times by 15 to 80 ms at 95% confidence.
   - The posterior is **compatible with H**.

Inference from the posterior

1. Formulate a **hypothesis**:

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.

Inference from the posterior

1. Formulate a **hypothesis**:
   - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).

Inference from the posterior

1. Formulate a **hypothesis**:
   - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):

Inference from the posterior

1. Formulate a **hypothesis**:
   - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
   - Condition B 95% CI = [-80, -15] ms.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:
    - The posterior suggests that Condition B decreases reaction times by 15 to 80 ms at 95% confidence.

Inference from the posterior

1. Formulate a **hypothesis**:
    - H: Condition B decreases reaction times relative to Condition A *by 100 ms*.
2. Choose **model specification** (including priors).
3. Collect data.
4. Calculate the **posterior** (fit the model):
    - Condition B 95% CI = [-80, -15] ms.
5. **Inference**:
    - The posterior suggests that Condition B decreases reaction times by 15 to 80 ms at 95% confidence.
    - The posterior is **not compatible with H**.

Inference with a ROPE

H0 vs H1

- H1 states that Condition B increases segment duration (alternative hypothesis), while H0 states that Condition B does not increase segment duration (null hypothesis, null effect).

Inference with a ROPE

H0 vs H1

- H1 states that Condition B increases segment duration (alternative hypothesis), while H0 states that Condition B does not increase segment duration (null hypothesis, null effect).
  - $H_1 : \beta > 0$

Inference with a ROPE

H0 vs H1

- H1 states that Condition B increases segment duration (alternative hypothesis), while H0 states that Condition B does not increase segment duration (null hypothesis, null effect).
  - $H_1 : \beta > 0$
  - $H_0 : \beta = 0$

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

- Define a region around $\beta = 0$ that practically corresponds to a null effect.

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

- Define a region around $\beta = 0$ that practically corresponds to a null effect.
  - For example: [-5, +5] ms ($-5 \geq \beta \leq +5$ = null effect).

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

- Define a region around $\beta = 0$ that practically corresponds to a null effect.
    - For example: [-5, +5] ms ($-5 \geq \beta \leq +5$ = null effect).
    - This ROPE has a width of 10 ms.

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

- Define a region around $\beta = 0$ that practically corresponds to a null effect.
  - For example: [-5, +5] ms ($-5 \geq \beta \leq +5$ = null effect).
  - This ROPE has a width of 10 ms.
- Collect data until the 95% CI of β has a width equal to or smaller than the width of the ROPE.

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

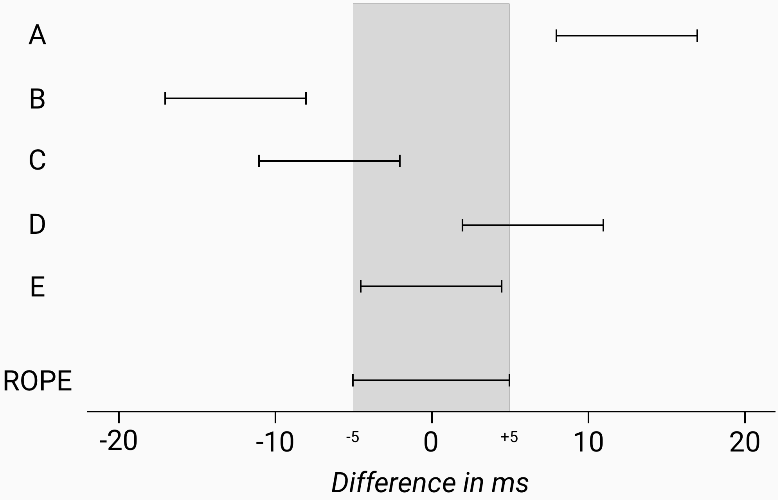
- Define a region around $\beta = 0$ that practically corresponds to a null effect.
  - For example: [-5, +5] ms ($-5 \geq \beta \leq +5$ = null effect).
  - This ROPE has a width of 10 ms.
- Collect data until the 95% CI of β has a width equal to or smaller than the width of the ROPE.
  - Choose a minimal sample size (ideally based on a prospective power analysis).

Inference with a ROPE

Region Of Practical Equivalence (ROPE):

- Define a region around $\beta = 0$ that practically corresponds to a null effect.
  - For example: [-5, +5] ms ($-5 \geq \beta \leq +5$ = null effect).
  - This ROPE has a width of 10 ms.
- Collect data until the 95% CI of β has a width equal to or smaller than the width of the ROPE.
  - Choose a minimal sample size (ideally based on a prospective power analysis).
  - Collect data and check 95% CI. If the width is greater than the ROPE, collect more data and repeat (*sequential testing*).

Inference with a ROPE



Inference with a ROPE

- We focus on the estimate **precision** of β.

Inference with a ROPE

- We focus on the estimate **precision** of β.
- Independent from the value of β.

Inference with a ROPE

- We focus on the estimate **precision** of β.
- Independent from the value of β.
- Higher precision means greater confidence in the estimated value of β.

Bayes Factor

The Bayes factor is the ratio of the likelihood of H1 to the likelihood of H2.

$$BF_{12} = \mathcal{L}(H_1)/\mathcal{L}(H_2)$$

Bayes Factor

BF	p(M1|D)	evidence
1–3	0.5–0.75	weak
3–20	0.75–0.95	positive
20–150	0.95–0.99	strong
> 150	> 0.99	very strong

Bayes Factor

```r
priors <- c(
  prior(normal(0, 500), class = Intercept),
  prior(cauchy(0, 15), class = sigma),
  prior(normal(0, 750), class = b, coef = "vowelo"),
  prior(normal(0, 750), class = b, coef = "vowelu"),
  prior(cauchy(0, 15), class = sd),
  prior(lkj(2), class = cor)
)
```

Bayes Factor

```
f1_3_bf <- brm(
  f1 ~ 1 + vowel + (1 + vowel | speaker),
  family = gaussian(),
  prior = priors,
  data = f_end,
  chains = 4,
  iter = 2000,
  file = "./cache/f1_3_bf",
  save_all_pars = TRUE
)
```

Bayes Factor

```r
priors <- c(
  prior(normal(0, 500), class = Intercept),
  prior(cauchy(0, 15), class = sigma),
  prior(normal(0, 750), class = b, coef = "vowelo"),
  prior(normal(0, 750), class = b, coef = "vowelu"),
  prior(normal(0, 750), class = b, coef = "c2_placevelar"),
  prior(cauchy(0, 15), class = sd),
  prior(lkj(2), class = cor)
)
```

Bayes Factor

```
f1_3_place <- brm(
  f1 ~ 1 + vowel + c2_place + (1 + vowel + c2_place | sp
  family = gaussian(),
  prior = priors,
  data = f_end,
  chains = 4,
  iter = 2000,
  file = "./cache/f1_3_place",
  save_all_pars = TRUE
)
```

Bayes Factor

```r
bf <- bayes_factor(f1_3_bf, f1_3_place)
```

```
## Iteration: 1
## Iteration: 2
## Iteration: 3
## Iteration: 4
## Iteration: 5
## Iteration: 6
## Iteration: 7
## Iteration: 8
## Iteration: 9
## Iteration: 1
## Iteration: 2
## Iteration: 3
```

Bayes Factor

```
bf
```

```
## Estimated Bayes factor in favor of f1_3_bf over f1_3_place: 0.58852
```

Bayes Factor

- BFs are ***VERY*** sensitive to priors and data.

Bayes Factor

- BFs are ***VERY*** sensitive to priors and data.
- Always calculate and report BFs by comparing models with increasingly narrower priors (at least 3-4).

Bayes Factor

- BFs are ***VERY*** sensitive to priors and data.
- Always calculate and report BFs by comparing models with increasingly narrower priors (at least 3-4).
- It's important to run *sensitivity analyses* that assess the influece of the priors on the posterior (not only if you you BFs, but always).

THE END