

Bayesian Model Comparison in R

Part I

Julia Haaf

Slides and Material

You can find the slides (and additional materials) here: <https://github.com/jstbcs/ws-bayesian-stats-r>.

Who Am I?

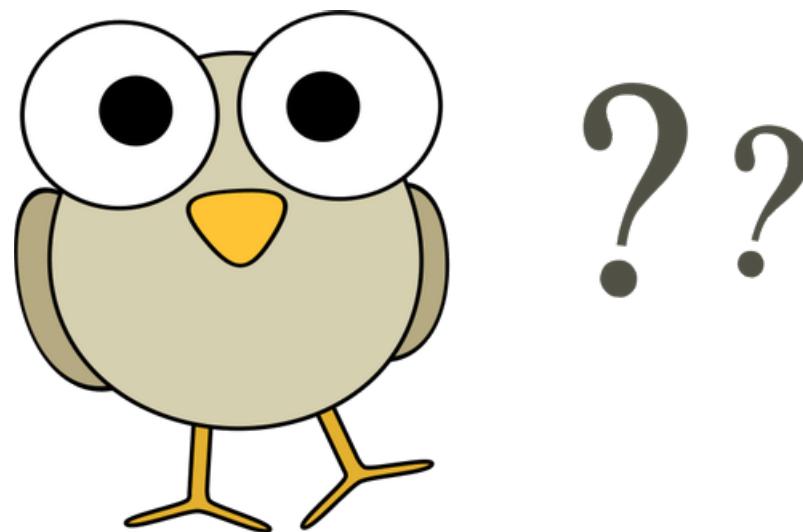
Julia Haaf

Assistant Professor at the Psychological Methods Department, University of Amsterdam

email: j.m.haaf@uva.nl

Job alert: Two temporary assistant professor positions at the Psychological Methods Unit (<https://www.academictransfer.com/nl/297540/assistant-professor-of-psychological-methods/>).

Who Are You?



Outline

1. Introduction to Bayesian inference
2. The BayesFactor package
3. *t*-Test
4. ANOVA
5. Linear model setup

Introduction to Bayesian inference

Example

- Effect of media stories on attitudes towards Syrian refugees.
- IV: Sympathetic media story about refugees or control story about food trends.
- DV: Participants' favorability ratings towards refugees.
- Also recorded: Political affiliation of the participants (conservative or liberal).



Example from Rouder, Haaf, & Aust (2018)

Usual approach

- Conduct a t -test to assess the effect of media stories on attitudes towards Syrian refugees.
- Conduct a 2x2 ANOVA to assess the effects of both media stories and political affiliation as well as their interaction.
- Not possible to provide evidence against these effects using classical statistics.
- Neither necessarily match the research questions associated with the study.
- Bayesian model comparison allows to better match inference and theoretical questions.
- But you can also do the regular stuff.

Three main theoretical positions

1. The plight of refugees generates an empathic response which results in an increase in favorability for conservative and liberals alike.
2. Confirmation bias (Tversky & Kahneman, 1974): People ignore media stories incompatible with their world view. Therefore, only liberals are affected.
3. Belief polarization (Cook & Lewandowsky, 2016): Liberals see refugees more favorably after viewing positive stories while conservatives might actually see them more negatively.

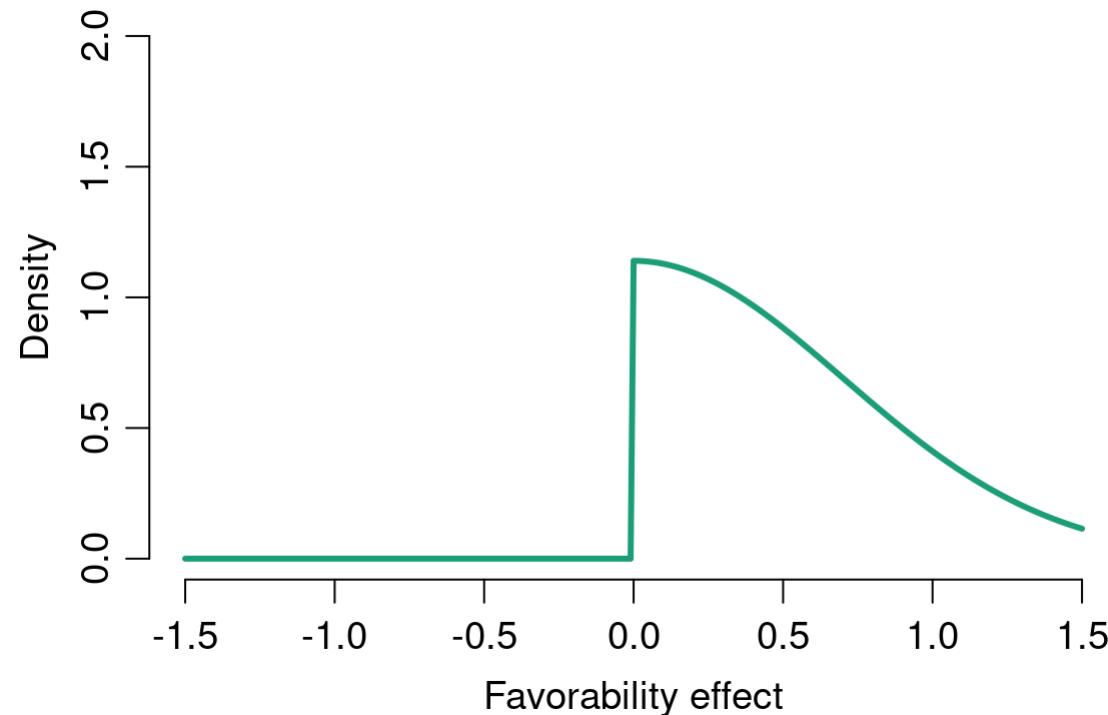


Testing position 1

- We need models

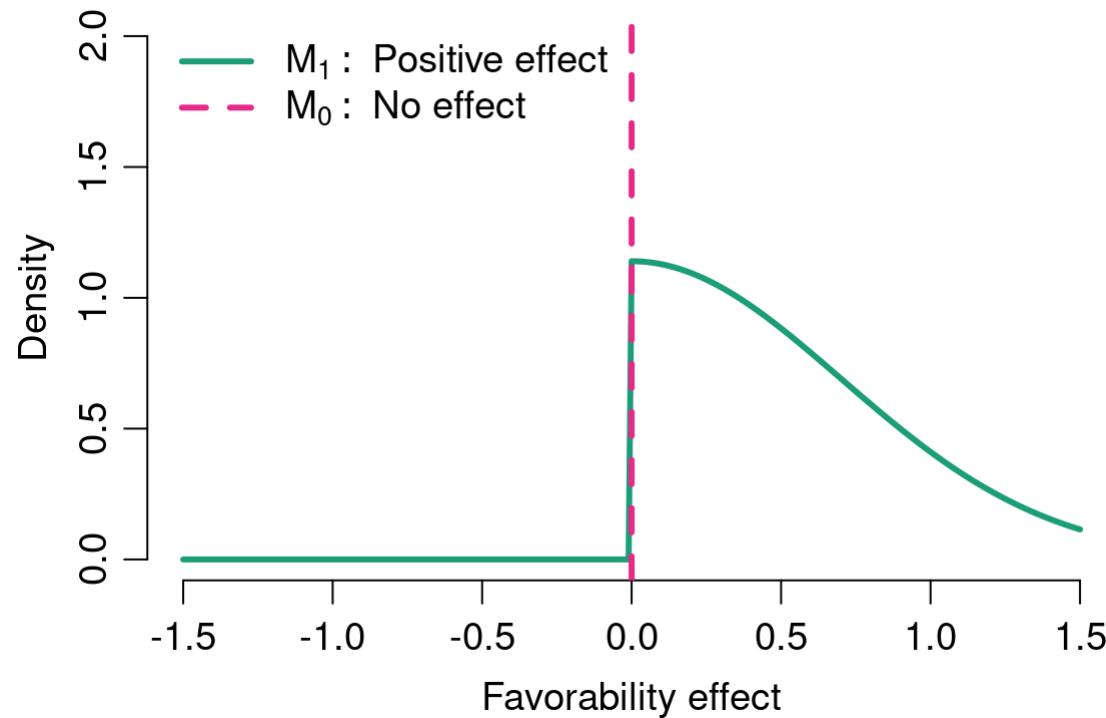
Testing position 1

- We need models
- \mathcal{M}_1 : The effect of media stories on attitudes is positive.



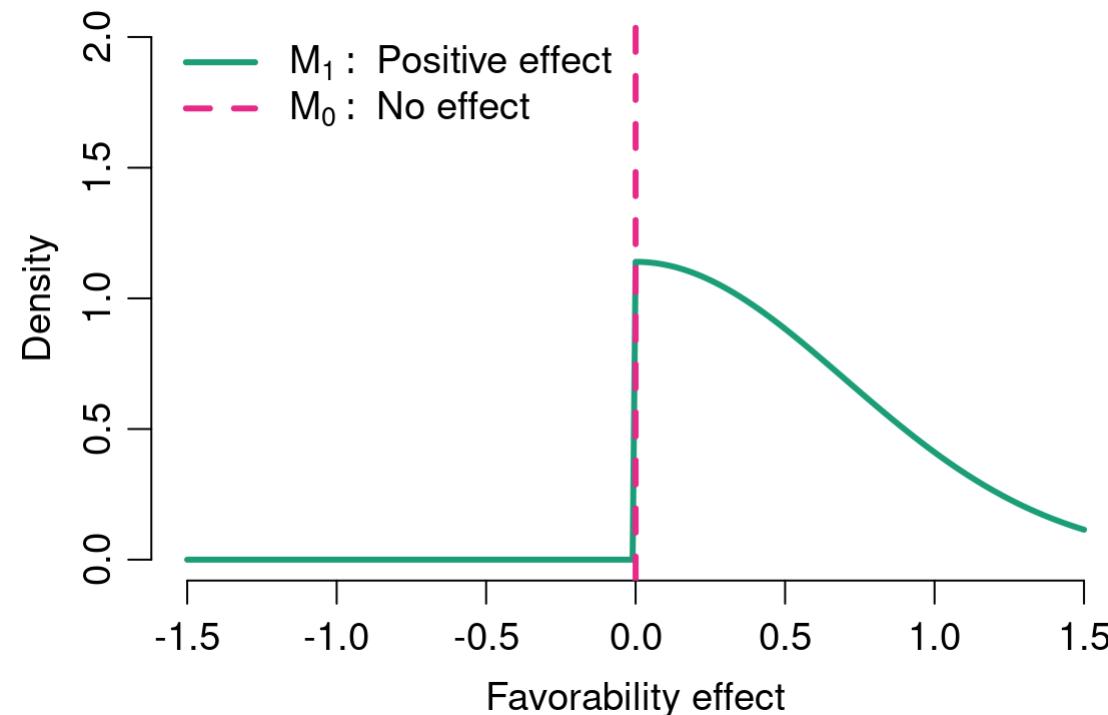
Testing position 1

- We need models
- \mathcal{M}_1 : The effect of media stories on attitudes is positive.
- \mathcal{M}_0 : There is no effect of media stories on attitudes.



Testing position 1

- The prior distribution indicates the plausibility of effects before the data have been collected.
- Because we make this fine-grained specification, we can also make predictions *before* collecting any data!



Testing position 1

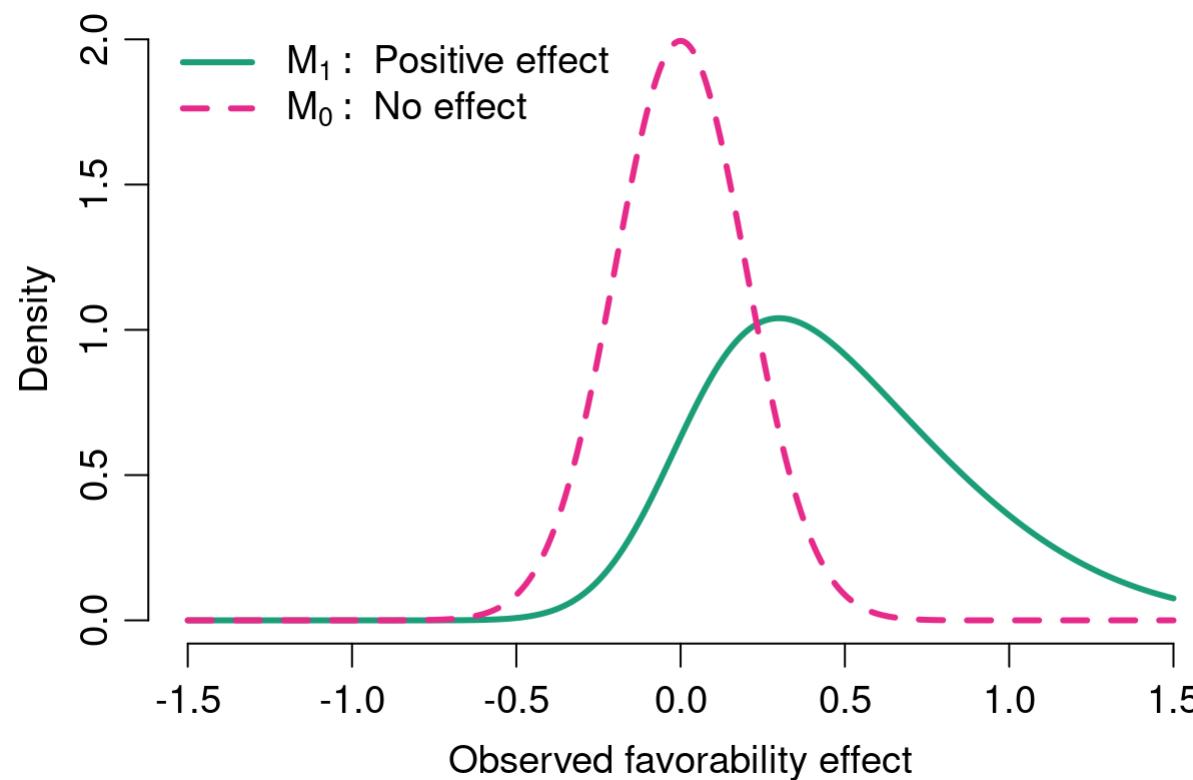
Evidence through predictions

- In Bayesian statistics, evidence for models reflects how well they predict data.
- Prior is on true values.
- Predictions are on data.

Testing position 1

Evidence through predictions

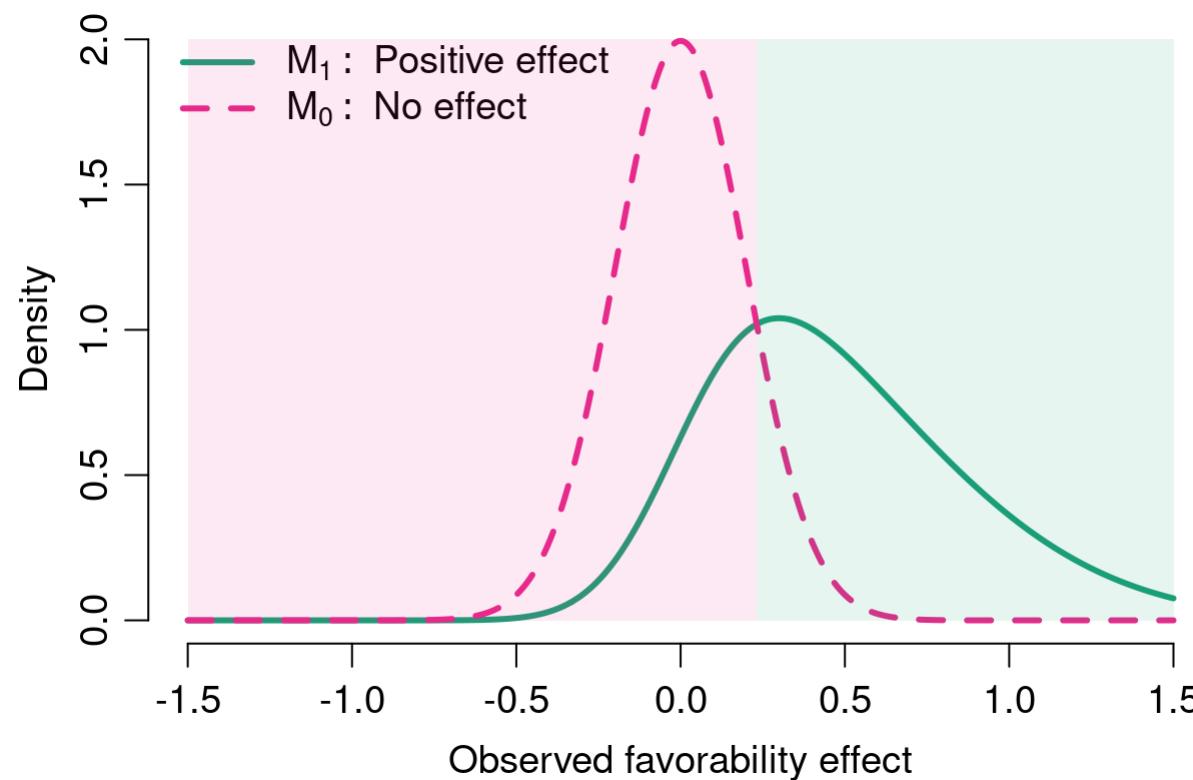
- In Bayesian statistics, evidence for models reflects how well they predict data.



Testing position 1

Evidence through predictions

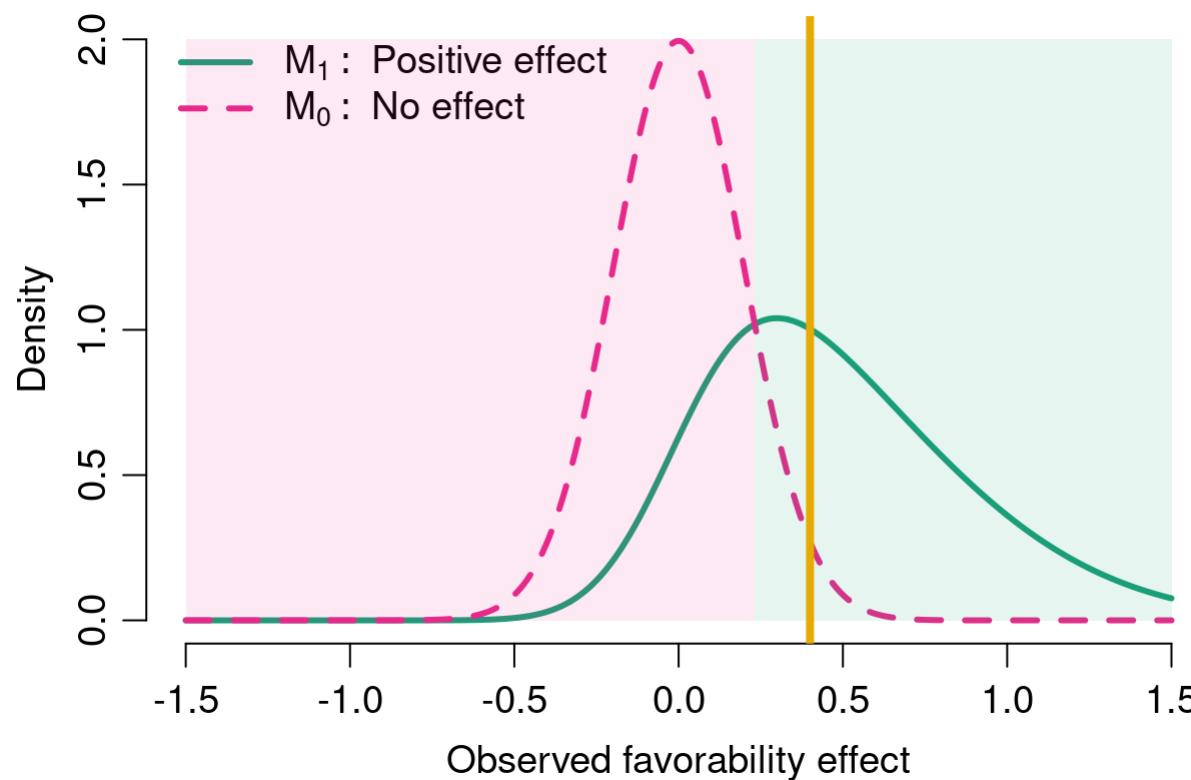
- In Bayesian statistics, evidence for models reflects how well they predict data.



Testing position 1

Evidence through predictions

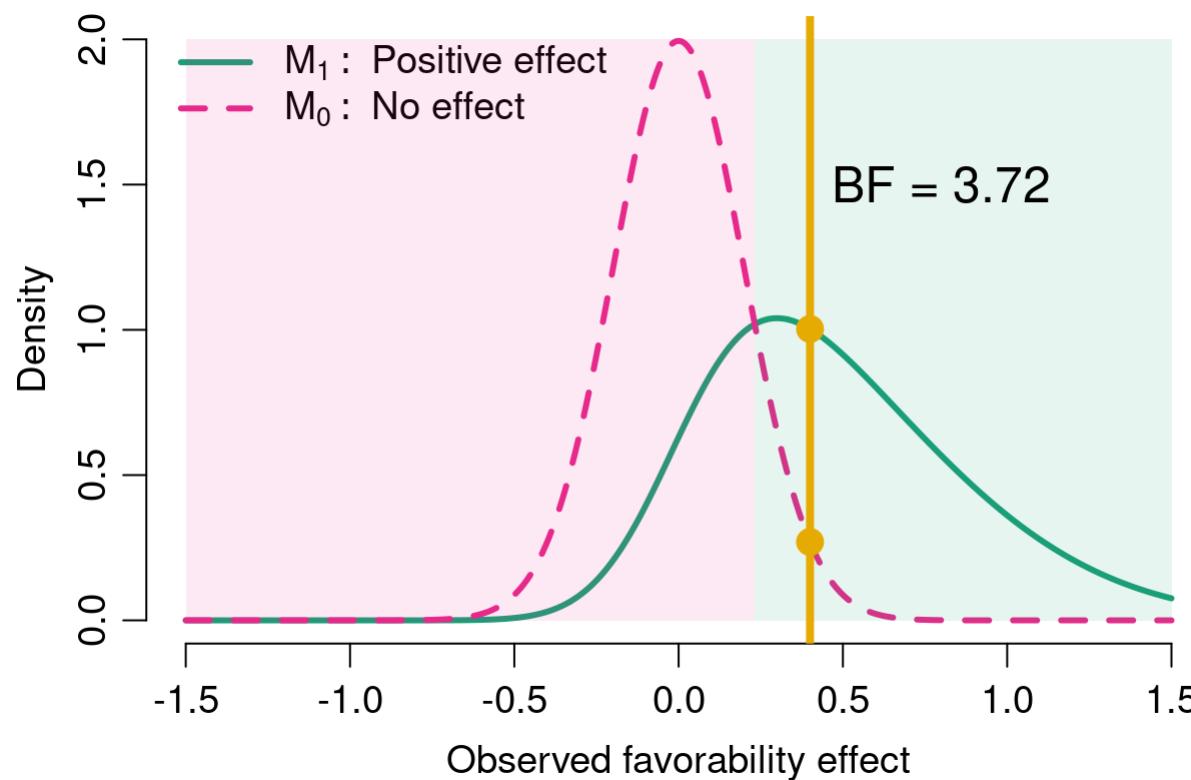
- In Bayesian statistics, evidence for models reflects how well they predict data.



Testing position 1

Evidence through predictions

- In Bayesian statistics, evidence for models reflects how well they predict data.



Bayesian model comparison in practice

The BayesFactor package

- Based on Rouder, Speckman, Sun, Morey, & Iverson (2009) (t -test), Rouder et al. (2012) (ANOVA), and Rouder & Morey (2012) (regression), total of > 4000 citations.
- Same setup is implemented in JASP and SPSS.
- Provides a convenient interface for commonly used models.
- Bayesian versions of typical frequentist tests using Bayes factor.



The BayesFactor package

Functionality

- General linear models (including linear mixed effects models): `generalTestBF`, `lmBF`
- Linear regression: `regressionBF`, `lmBF`
- Linear correlation: `correlationBF`
- *t*-tests: `ttestBF`
- Meta-analytic *t*-tests: `meta.ttestBF`
- ANOVA: `anovaBF`, `lmBF`
- Contingency tables: `contingencyTableBF`
- Single proportions: `proportionBF`

The BayesFactor package

Functionality

- **General linear models** (including linear mixed effects models): `generalTestBF`, `lmBF`
- Linear regression: `regressionBF`, `lmBF`
- Linear correlation: `correlationBF`
- ***t*-tests:** `ttestBF`
- Meta-analytic *t*-tests: `meta.ttestBF`
- **ANOVA:** `anovaBF`, `lmBF`
- Contingency tables: `contingencyTableBF`
- Single proportions: `proportionBF`

***t*-Test**

t-Test using BayesFactor

Can be used to test our position 1: The media story increases favorability ratings of refugees.

```
BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study)
```

If you want to understand formulas in R better try this [tutorial](#) (for general setup) and this [tutorial](#) (for mixed models).

Prior settings

- \mathcal{M}_1 : The effect of media stories on attitudes is positive.
- \mathcal{M}_0 : There is no effect of media stories on attitudes.
- Attitudes might be normally distributed: $Y_{ij} \sim \text{Normal}(\mu + x_j\theta, \sigma^2)$.
- For model \mathcal{M}_0 $\theta = 0$.
- For model \mathcal{M}_1 we need a prior distribution on the effect, θ
- Prior on σ^2 and μ ?

Four reasons for using informed priors

1. It is necessary
2. You gotta know something

2. You gotta know something

You are...

- A researcher in psychology.
- Working in their field of expertise.
- Standing on the shoulders of previous generations of researchers.

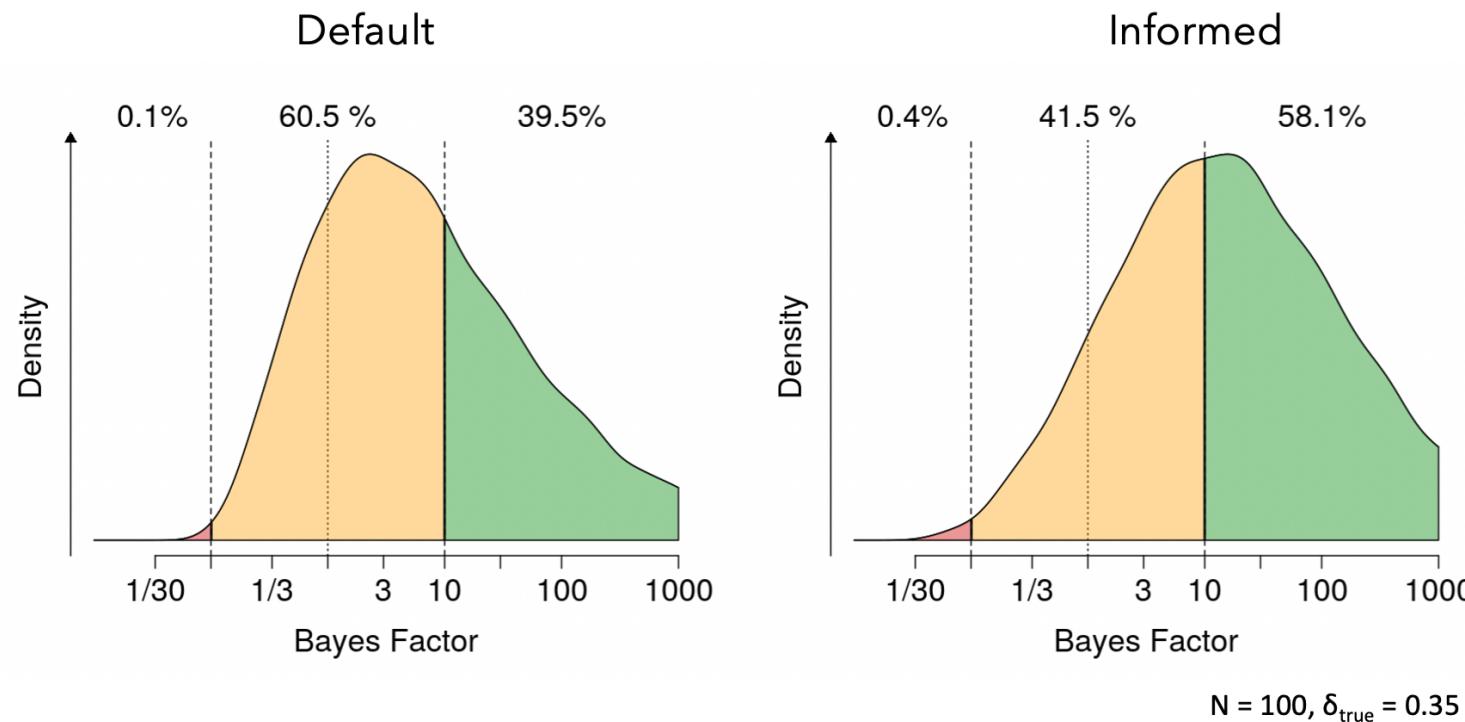


Four reasons for using informed priors

1. It is necessary
2. You gotta know something
3. Informed priors → informative tests

3. Informed priors → informative tests

Obtain strong evidence by increasing model discriminability.



Stefan, Gronau, Schönbrodt, & Wagenmakers (2019)

Four reasons for using informed priors

1. It is necessary
2. You gotta know something
3. Informed priors → informative tests
4. Informed priors ensure identifiability

4. Informed priors ensure identifiability

- Complex models often suffer from identifiability issues.

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =  
## control$checkConv, : Model failed to converge with max|grad| = 1.52673  
## (tol = 0.001, component 17)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable: very large ei  
genvalue  
## - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio  
## - Rescale variables?
```

- Informed priors introduce probabilistic constraints.

The blessing of default priors

- BayesFactor package in R and JASP use default priors for the t-test called JZS prior.
- Prior structure is consistent.
- Joint prior on μ and σ^2 is a Jeffreys prior.
- User-defined informed prior settings are on the scale of the effect size:
- $\delta = \theta/\sigma$ (think Cohen's d).

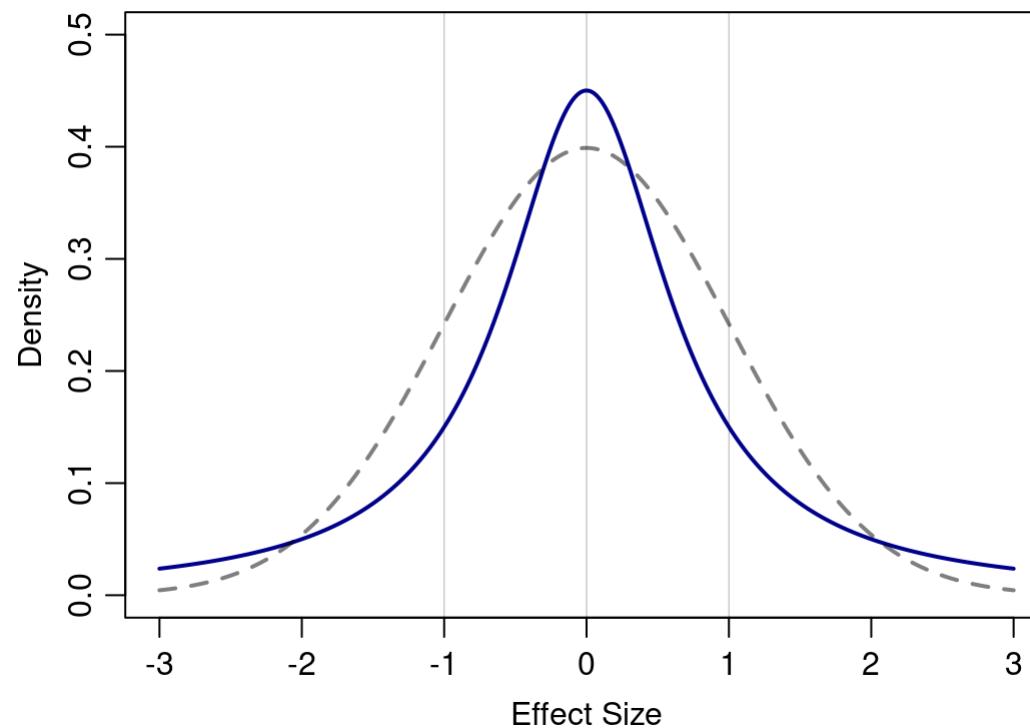
The blessing of default priors

- BayesFactor package in R and JASP use default priors for the t-test called JZS prior.
- Prior structure is consistent.
- User-defined prior settings are on the scale of the effect size.

```
BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study  
, rscale = 1 / sqrt(2))
```

The blessing of default priors

```
BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study  
, rscale = 1 / sqrt(2))
```



The blessing of default priors

Figure code

```
par(mar = c(3,3,.5,.5), mgp = c(2,.7,0))
x <- seq(-3, 3, .01)
y <- dcauchy(x, 0, 1/sqrt(2))
plot(x, y, type = "l", lwd = 2
      , ylab = "Density", xlab = "Effect Size"
      , col = "darkblue", ylim = c(0, .5))
lines(x, dnorm(x), col = adjustcolor(1, .5), lty = 2, lwd = 2)
abline(v = c(-1, 0, 1), lwd = 0.5, col = adjustcolor(1, alpha.f = 0.3))
```

Prior prediction

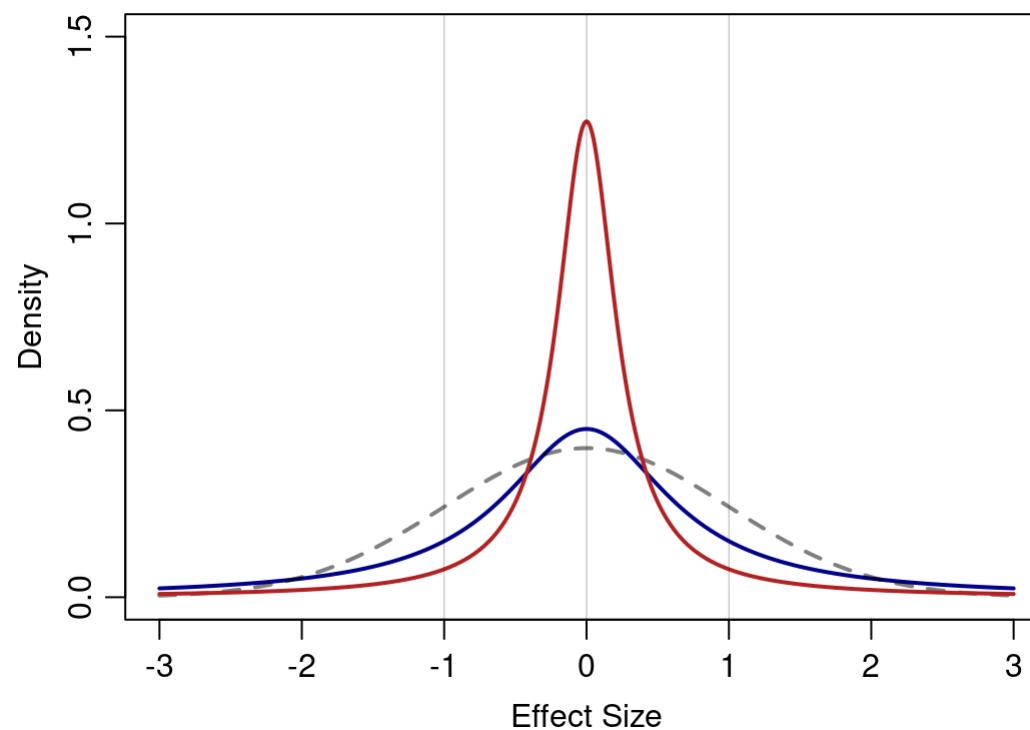
- Prior → we can make predictions on data.
- This is a bit tricky with the JZS prior though...

Prior prediction

```
sim.pred.fun <- function(scale){  
  delta <- rcauchy(1, location = 0, scale = scale)  
  mu <- runif(1, -1000, 1000)  
  s2 <- MCMCpack::rinvgamma(1, 0.5, 0.5)  
  
  I <- 100  
  cond <- rep(c(-1/sqrt(2), 1/sqrt(2)), each = I)  
  y <- rnorm(2 * I, mu + cond * delta * sqrt(s2), sqrt(s2))  
  
  return(diff(tapply(y, cond, mean)))  
}  
  
res <- replicate(100000, sim.pred.fun(1/sqrt(2)))  
round(quantile(res), 2)  
  
##          0%        25%        50%        75%       100%  
## -1179562.46     -1.74     -0.01      1.70    155159.42
```

The blessing of default priors

```
BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study  
, rscale = 1 / 4)
```



Your turn!

1. Plot the default prior on effect size.
2. Adjust the scale until it matches your expectations of the outcome of the study.
3. Conduct a t -test using your setting.
4. Do you think any two people in the workshop get the same result?

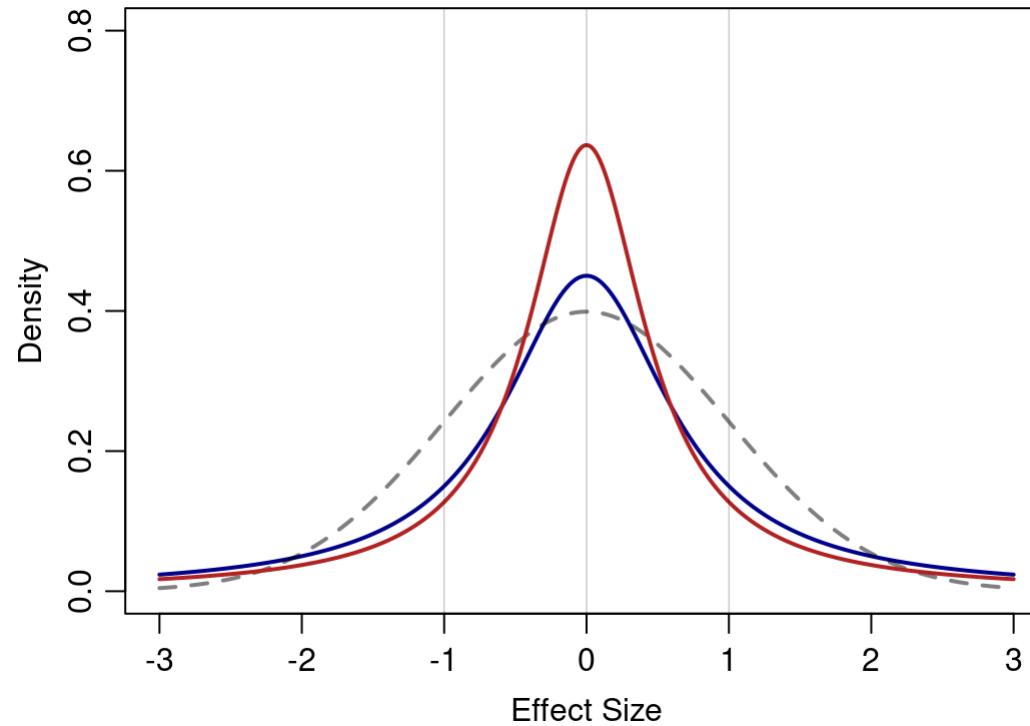


What was your scale setting?

Put your setting in the zoom chat!



What was your scale setting?



t-Test using BayesFactor

Output

```
BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study  
, rscale = 1 / 2)  
  
## Bayes factor analysis  
## -----  
## [1] Alt., r=0.5 : 2.142883 ±0%  
##  
## Against denominator:  
##   Null, mu1-mu2 = 0  
## ---  
## Bayes factor type: BFindepSample, JZS
```

Back to media stories about refugees

The plight of refugees generates an empathic response which results in an increase in favorability for conservative and liberals alike.

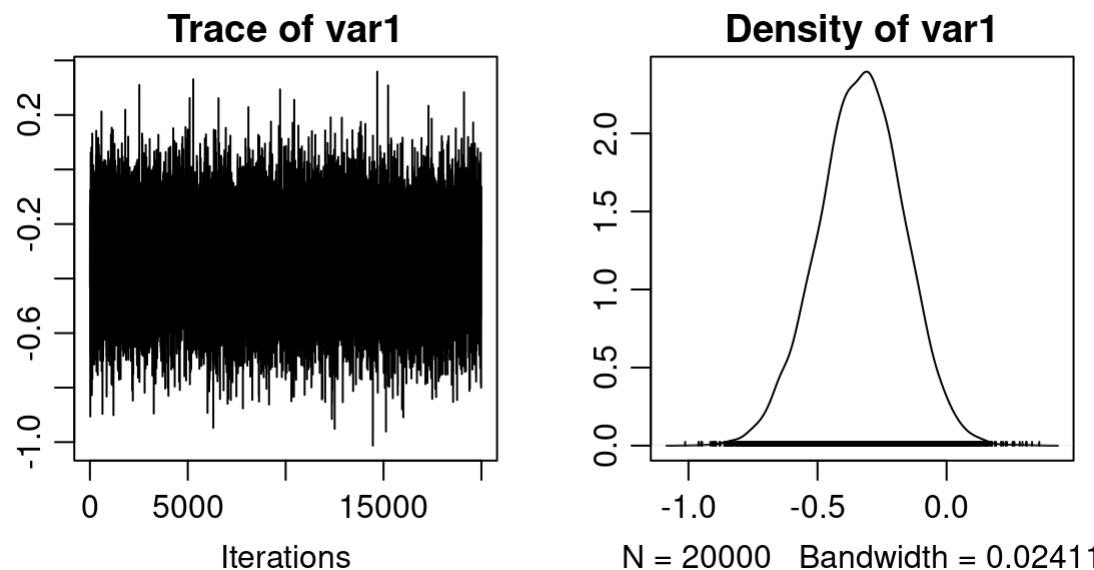
- Does the t -test as we used it address this statement?

Estimating the effect

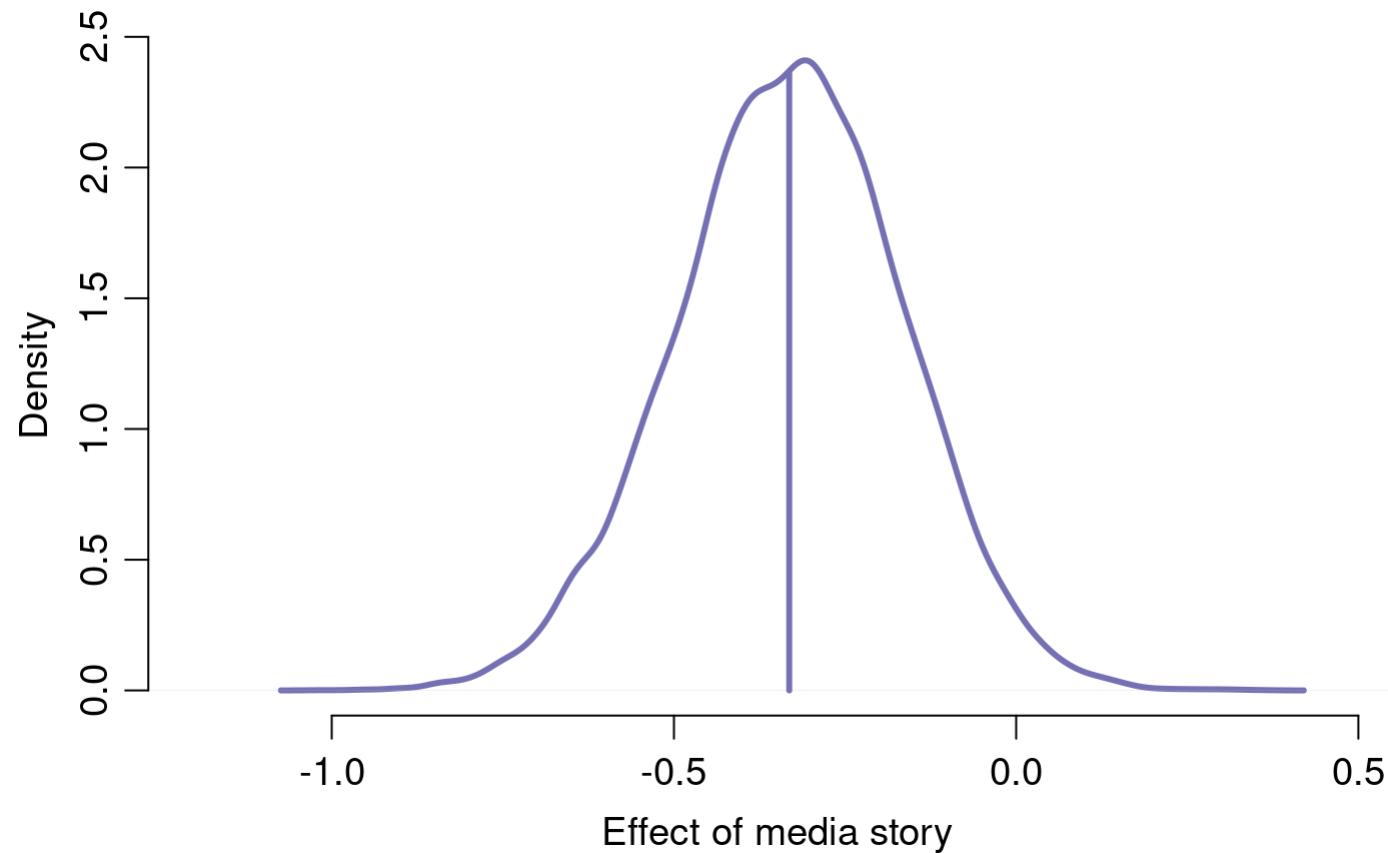
```
est.attitude <- BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study, rscale = 1 / 2  
, posterior = T, iterations = 20000)
```

Estimating the effect

```
est.attitude <- BayesFactor::ttestBF(formula = attitude ~ story  
, data = media_study, rscale = 1 / 2  
, posterior = T, iterations = 20000)  
  
plot(est.attitude[, "beta (Neutral - Refugee plight)"])
```



Estimating the effect



Back to media stories about refugees

The plight of refugees generates an empathic response which results in an increase in favorability for conservative and liberals alike.

- Does the t -test as we used it address this statement?
- We need an ordinal constraint.
- Ordinal constraint: Specification of the direction of an effect.

One-sided tests with BayesFactor

- One-sided t -tests are easy with BayesFactor.

One-sided tests with BayesFactor

- One-sided t -tests are easy with BayesFactor.

```
BayesFactor::ttestBF(formula = attitude ~ story  
                      , data = media_study  
                      , rscale = 1 / 2  
                      , nullInterval = c(-Inf, 0))  
  
## Bayes factor analysis  
## -----  
## [1] Alt., r=0.5 -Inf<d<0 : 4.196947 ±0%  
## [2] Alt., r=0.5 !(-Inf<d<0) : 0.08881904 ±0%  
##  
## Against denominator:  
##   Null, mu1-mu2 = 0  
## ---  
## Bayes factor type: BFindepSample, JZS
```

Your turn!

1. New data set, this time real data.
2. Run a Bayesian t-test, try out ordinal constraints and estimation, maybe even prior prediction.
3. Ask questions while I am here!

```
SourceURL <- "https://raw.githubusercontent.com/PerceptionCognitionLab/data0/master  
/contexteffects/FlankerStroopSimon/cleaning.R"  
devtools::source_url(SourceURL)
```

```
library(tidyverse)
```

```
stroop.agg <- stroop %>%  
  group_by(ID, congruency) %>%  
  summarize(mrt = mean(RT)) %>%  
  spread(congruency, mrt)
```



Analysis of variance

ANOVA

- Based on Rouder et al. (2012).
- Between-subject, within-subject and mixed designs.
- Random effects.
- Perhaps the most thought out functionality in BayesFactor.

ANOVA models

Considered

Not Considered

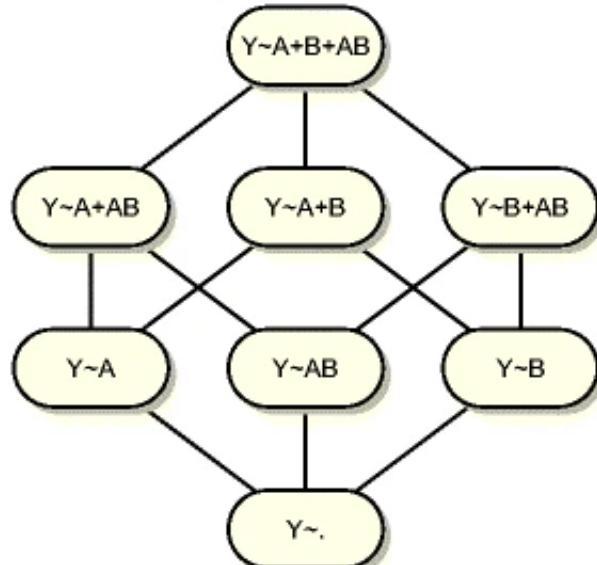
Implausible

A One-Factor

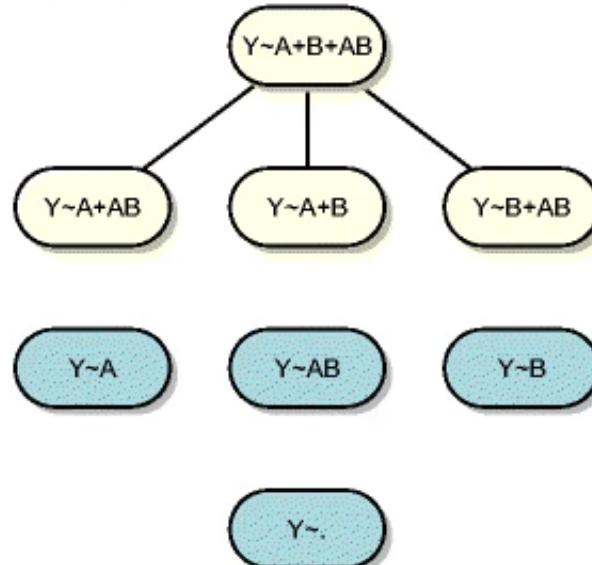
$Y \sim A$

$Y \sim .$

B Two Factors



C ANOVA



ANOVA parameterization

- Extension to t -test.
- Effects are parameterized as $\delta = \theta/\sigma$ (effect size).
- Coding is orthonormal (depends on the number of levels per factor).
- Most simple case is two levels: $x_1 = -1/\sqrt{2}, x_2 = 1/\sqrt{2}$.

ANOVA in practice

```
media_study$story <- factor(media_study$story)
media_study$political_affiliation <- factor(media_study$political_affiliation)

BayesFactor::anovaBF(formula = attitude ~ story * political_affiliation
, data = media_study)
```

ANOVA in practice

rscaleEffects = A named vector of prior settings for individual factors, overriding **rscaleFixed** and **rscaleRandom**. Values are scales, names are factor names.

```
BayesFactor::anovaBF(formula = attitude ~ story * political_affiliation  
, data = media_study  
, rscaleEffects = c("story" = 1/2  
, "political_affiliation" = 1/2  
, "story:political_affiliation" = 1/3))
```

ANOVA in practice

Output

```
BayesFactor::anovaBF(formula = attitude ~ story * political_affiliation  
, data = media_study  
, rscaleEffects = c("story" = 1/2  
, "political_affiliation" = 1/2  
, "story:political_affiliation" = 1/3))  
  
## Bayes factor analysis  
## -----  
## [1] story : 1.835452 ±0%  
## [2] political_affiliation : 191.5876 ±0%  
## [3] story + political_affiliation : 468.3391 ±0.73%  
## [4] story + political_affiliation + story:political_affiliation : 1156.833 ±7.65%  
##  
## Against denominator:  
##   Intercept only  
## ---  
## Bayes factor type: BFlinearModel, JZS
```

ANOVA models

Considered

Not Considered

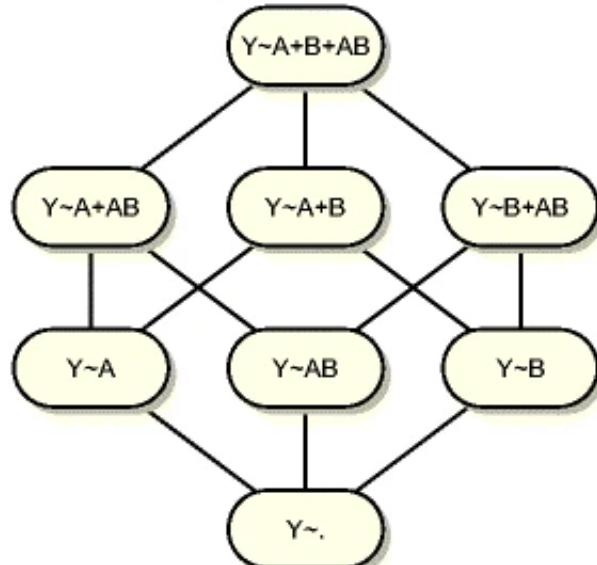
Implausible

A One-Factor

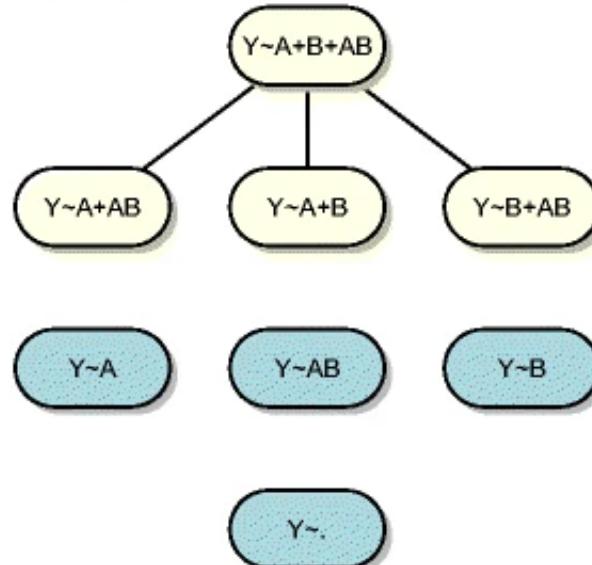
$Y \sim A$

$Y \sim .$

B Two Factors



C ANOVA

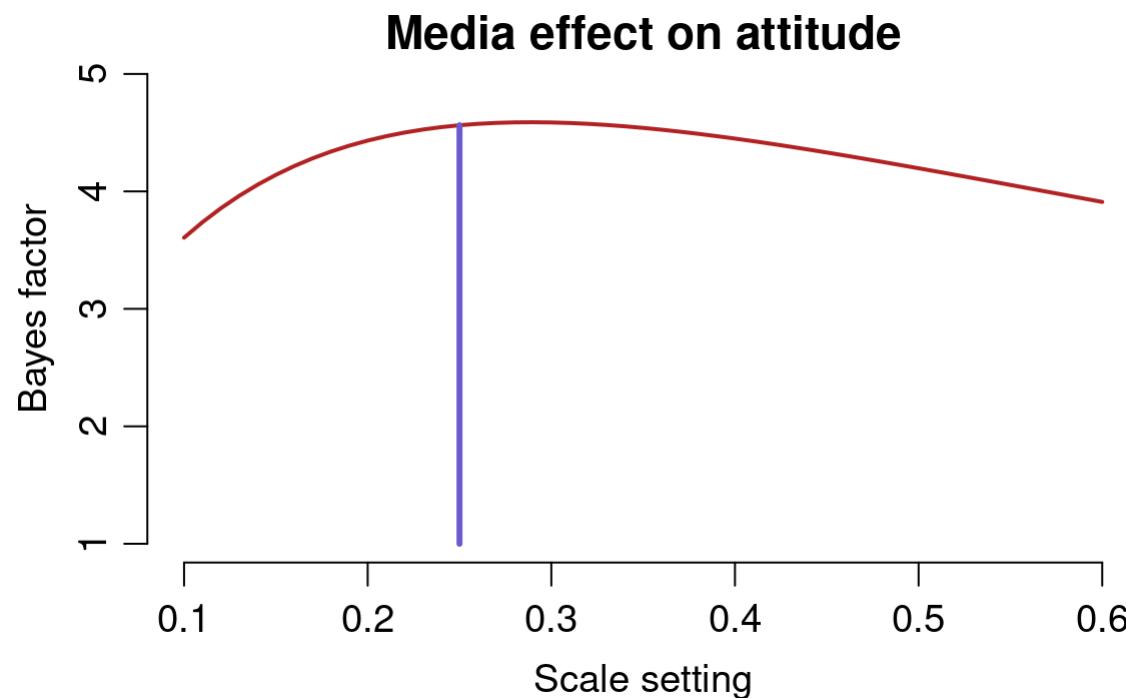


Prior Sensitivity

- How does your prior influence the results of the analysis?
- Redo the analysis for a range of priors.
- If the results are (relatively) stable then we may trust them more.

Prior Sensitivity

- How does your prior influence the results of the analysis?
- Redo the analysis for a **reasonable** range of priors.
- If the results are (relatively) stable then we may trust them more.



Prior Sensitivity

For two scales

Explore the range of settings using all combinations.

a

0.3

0.5

0.7

0.3

0.5

0.7

0.3

0.5

b

0.3

0.3

0.3

0.5

0.5

0.5

0.7

Sensitivity analysis versus multiverse analysis

- In a multiverse analysis you draw your conclusions based on the entire space of possible parameter settings.
- The sensitivity analysis provides information about the robustness of your analysis and conclusions, it is not meant to replace the main analysis.
- Remember, you set the initial prior with care.



Your turn!

1. Run the ANOVA analysis on the `dat_media.csv`.
2. Conduct a sensitivity analysis for all three prior scales using *reasonable ranges* of settings.
3. Is the variability in Bayes factor reasonable?

```
geturl <- "https://raw.githubusercontent.com/jstbcs/ws-bayesian-stats-r/  
main/data/dat_media.csv"  
media_study <- read.csv(geturl)
```



Ordinal constraints

- Not as straight-forward as with the t -test.
- Not as impossible as with frequentist ANOVA.

Back to the three main theoretical positions

1. The plight of refugees generates an empathic response which results in an increase in favorability for conservative and liberals alike.
2. Confirmation bias (Tversky & Kahneman, 1974): People ignore media stories incompatible with their world view. Therefore, only liberals are affected.
3. Belief polarization (Cook & Lewandowsky, 2016): Liberals see refugees more favorably after viewing positive stories while conservatives might actually see them more negatively.



Back to the three main theoretical positions

1. The plight of refugees generates an empathic response which results in an increase in favorability for conservative and liberals alike.
2. Confirmation bias (Tversky & Kahneman, 1974): People ignore media stories incompatible with their world view. Therefore, only liberals are affected.
3. **Belief polarization** (Cook & Lewandowsky, 2016): Liberals see refugees more favorably after viewing positive stories while conservatives might actually see them more negatively.



Testing position 3

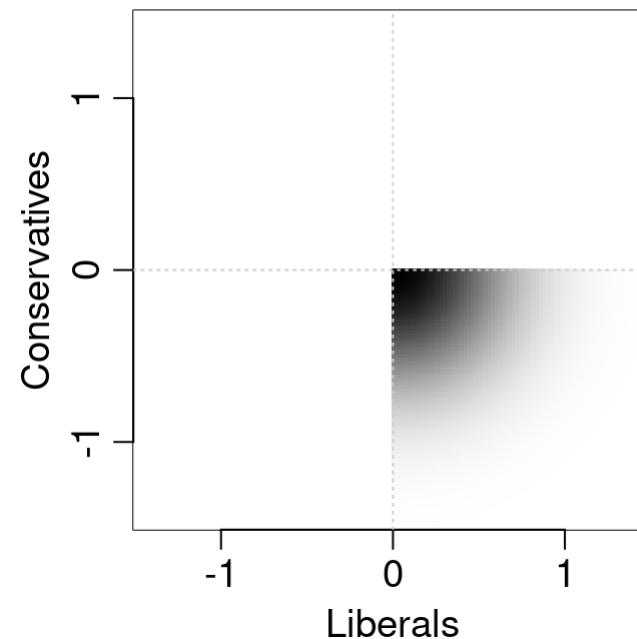
1. **Belief polarization** (Cook & Lewandowsky, 2016): Liberals see refugees more favorably after viewing positive stories while conservatives might actually see them more negatively.

\mathcal{M}_2 : The effect of the refugee media story is positive for liberals and negative for conservatives.

Testing position 3

1. **Belief polarization** (Cook & Lewandowsky, 2016): Liberals see refugees more favorably after viewing positive stories while conservatives might actually see them more negatively.

\mathcal{M}_2 : The effect of the refugee media story is positive for liberals and negative for conservatives.



Testing position 3

Encompassing approach (Klugkist, Laudy, & Hoijtink, 2005):

1. Calculate/estimate the prior probability $P(\text{constraint})$ for the ordinal constraint.
2. Estimate the posterior probability $P(\text{constraint} | Y)$ from posterior samples of the model.
3. The Bayes factor between the unconstrained and the ordinal-constrained model can be estimated as

$$BF = \frac{P(\text{constraint})}{P(\text{constraint} | Y)}$$

Testing position 3

Prior probability

- Two constraints, symmetric priors: $0.5 \times 0.5 = 0.25$.
- Can also be estimated:

$$Y_{RL} = \mu + \theta_1 + \theta_2 + \theta_3$$

$$Y_{CL} = \mu - \theta_1 + \theta_2 - \theta_3$$

$$Y_{RC} = \mu + \theta_1 - \theta_2 - \theta_3$$

$$Y_{CC} = \mu - \theta_1 - \theta_2 + \theta_3$$

- Constraints refer to $Y_{RL} - Y_{CL}$ and $Y_{RC} - Y_{CC}$, respectively.

- $$Y_{RL} - Y_{CL} = 2\theta_1 + 2\theta_3$$

$$Y_{RC} - Y_{CC} = 2\theta_1 - 2\theta_3$$

Testing position 3

Prior probability

$$Y_{RL} - Y_{CL} = 2\theta_1 + 2\theta_3$$
$$Y_{RC} - Y_{CC} = 2\theta_1 - 2\theta_3$$

```
M <- 100000
theta1 <- rcauchy(M)
theta3 <- rcauchy(M)

mean(theta1 + theta3 > 0 & theta1 - theta3 < 0)

## [1] 0.24943
```

Testing position 3

Posterior probability

```
tmp <- BayesFactor::anovaBF(formula = attitude ~ story * political_affiliation  
    , data = media_study  
    , rscaleEffects = c("story" = 1/2  
                      , "political_affiliation" = 1/2  
                      , "story:political_affiliation" = 1/3))  
  
est <- BayesFactor::posterior(model = tmp  
    , iterations = 20000  
    , index = 4)
```

Testing position 3

Posterior probability

Reconstructing cell means and effects from the estimates.

```
Y_cl <- est[, "story-Neutral"] +  
  est[, "story:political_affiliation-Neutral.&.Liberals"]  
  
Y_rl <- est[, "story-Refugee plight"] +  
  est[, "story:political_affiliation-Refugee plight.&.Liberals"]  
  
Y_cc <- est[, "story-Neutral"] +  
  est[, "story:political_affiliation-Neutral.&.Conservatives"]  
  
Y_rc <- est[, "story-Refugee plight"] +  
  est[, "story:political_affiliation-Refugee plight.&.Conservatives"]  
  
(postprob <- mean(Y_rl - Y_cl > 0 & Y_rc - Y_cc < 0))  
## [1] 0.36585
```

Testing position 3

Posterior probability

Bayes factor:

```
(BF_u3 <- 0.25 / postprob)
```

```
## [1] 0.6833402
```

Testing position 3

Posterior probability

```
bftab <- tmp@bayesFactor[, 1:2]
bftab[, 1] <- exp(bftab[, 1])
bftab[5, 1] <- bftab[4, 1] / BF_u3
rownames(bftab)[5] <- "Position 3"
knitr::kable(bftab)
```

| | bf | error |
|---|-------------|-----------|
| story | 1.835452 | 0.0000000 |
| political_affiliation | 191.587618 | 0.0000000 |
| story + political_affiliation | 464.094900 | 0.0202599 |
| story + political_affiliation + story:political_affiliation | 1058.549700 | 0.0086121 |
| Position 3 | 1549.081631 | NA |

Your turn!

- Go through the code and make sure you understand what is going on.
- Say we wanted to test whether the effect of media story is positive both for liberals and conservatives. What would change and what would stay the same?
- Implement the new constraint and report the Bayes factor for the new position 4 vs. the intercept-only model.



All the models

All the models

- General linear models (including linear mixed effects models): `generalTestBF`, `lmBF`
- Linear regression: `regressionBF`, `lmBF`
- Linear correlation: `correlationBF`
- *t*-tests: `ttestBF`
- Meta-analytic *t*-tests: `meta.ttestBF`
- ANOVA: `anovaBF`, `lmBF`
- Contingency tables: `contingencyTableBF`
- Single proportions: `proportionBF`

All the models

generalTestBF

This function computes Bayes factors corresponding to restrictions on a full model.

```
data(puzzles)
BayesFactor::generalTestBF(RT ~ shape * color + ID
                           , data = puzzles
                           , whichRandom = "ID")

## Bayes factor analysis
## -----
## [1] shape          : 0.6114754 ±0.01%
## [2] color          : 0.6114754 ±0.01%
## [3] shape + color  : 0.3865341 ±0.82%
## [4] shape + color + shape:color : 0.1328641 ±1.55%
## [5] ID             : 111516.6  ±0%
## [6] shape + ID     : 316107.6  ±0.98%
## [7] color + ID     : 313479.8  ±1.41%
## [8] shape + color + ID : 1309516  ±1.67%
## [9] shape + color + shape:color + ID : 512538.7  ±1.99%
##
```

All the models

generalTestBF

This function computes Bayes factors corresponding to restrictions on a full model.

lmBF

This function computes Bayes factors, or samples from the posterior, of specific linear models (either ANOVA or regression).

```
BayesFactor::lmBF(RT ~ shape + color + shape:color + ID  
, data = puzzles  
, whichRandom = "ID")
```

```
## Bayes factor analysis  
## -----  
## [1] shape + color + shape:color + ID : 488891.2 ±3%  
##  
## Against denominator:  
##   Intercept only
```

Continuous predictors

Rouder & Morey (2012)

Details

`regressionBF` computes Bayes factors to test the hypothesis that slopes are 0 against the alternative that all slopes are nonzero.

The vector of observations y is assumed to be distributed as

$$y \sim \text{Normal}(\alpha \mathbf{1} + X\beta, \sigma^2 I).$$

The joint prior on α, σ^2 is proportional to $1/\sigma^2$, the prior on β is

$$\beta \sim \text{Normal}(0, N g \sigma^2 (X'X)^{-1}).$$

where $g \sim \text{InverseGamma}(1/2, r/2)$. See Liang et al. (2008) section 3 for details.

Your turn

- Now it is time to play.
- Repeat the analysis for `dat_media` using `generalTestBF` and `lmBF`.
- Alternatively, use your own data or continue with `bugs.csv`.



Bringing it all together

- `BayesFactor` can do a lot of conventional tests, but in Bayesian.
- Model comparison instead of parameter testing.
- For a full Bayesian analysis there are many tools in addition to Bayes factor:
- Prior prediction,
- Estimation,
- Sensitivity analysis,
- Data plotting(!), ...



Thank you!



Cook, J., & Lewandowsky, S. (2016). Rational irrationality: Modeling climate change belief polarization using Bayesian networks. *Topics in Cognitive Science*, 8(1), 160–179. Retrieved from <https://dx.doi.org/10.1111/tops.12186>

Klugkist, I., Laudy, O., & Hoijtink, H. (2005). Inequality constrained analysis of variance: A bayesian approach. *Psychological Methods*, 10(4), 477.

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in anova. *Psychonomic Bulletin & Review*, 23(6), 1779–1786. Retrieved from <https://doi.org/10.3758/s13423-016-1026-5>

Rouder, J. N., Haaf, J. M., & Aust, F. (2018). From theories to models to predictions: A Bayesian model comparison approach. *Communication Monographs*, 85, 41–56. Retrieved from <https://doi.org/10.1080/03637751.2017.1394581>

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research*, 47, 877–903. Retrieved from <http://dx.doi.org/10.1080/00273171.2012.734737>