Introduction to Bayesian Statistics with R

Day 2

Oliver Lindemann

Erasmus University Rotterdam, NL

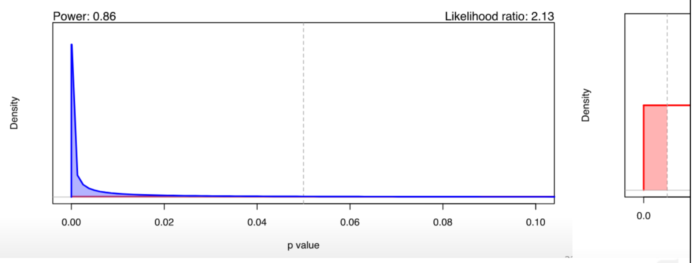
The Lindley "paradox"

How much evidence does the result $p = 0.05$ provide?

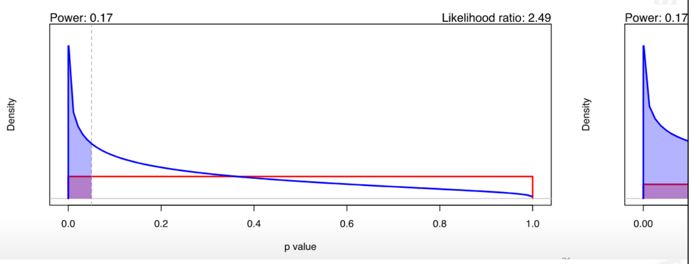
- Classical statistics has no definition of evidence
- The amount of evidence that $p = 0.05$ provides depends on the sample size, N

To begin, we consider the distribution of p values.

Distribution of p values under $\mathcal{H}_0$



Distribution of p values under $\mathcal{H}_1$



The Lindley "paradox"

This is considered a "paradoxical" result.

- $p = 0.05$ is considered "enough evidence" to reject $\mathcal{H}_0$
- However, with high enough power $p = 0.05$ becomes even evidence for $\mathcal{H}_0$!
- Why? With high N, p values should get closer and closer to 0, but the criterion $p < 0.05$ doesn't change with N.

$\Rightarrow$ We can not interpret p-values as amount of evidence and use it merely as a binary decision logic.

Bayes Hypothesis Testing

Two models: $\mathcal{H}_0$ and $\mathcal{H}_1$.

- Which model is better supported by the data?
- The model that predicted the data best!

The ratio of predictive performance is known as the **Bayes factor** (Jeffreys, 1961).

Bayes Factor

Comparing two Models, $\mathcal{H}_1, \mathcal{H}_0$

- $p(\mathcal{H}_1|D) = p(D|\mathcal{H}_1)\, p(\mathcal{H}_1) \,/\, p(D)$
- $p(\mathcal{H}_0|D) = p(D|\mathcal{H}_0)\, p(\mathcal{H}_0) \,/\, p(D)$

$$\frac{p(\mathcal{H}_1|D)}{p(\mathcal{H}_0|D)} = \frac{p(D|\mathcal{H}_1)}{p(D|\mathcal{H}_0)} \cdot \frac{p(\mathcal{H}_1)}{p(\mathcal{H}_0)}$$

Bayes Factor: Comparing models: $\mathcal{H}_1$ and $\mathcal{H}_0$

$$\underbrace{\frac{p(\mathcal{H}_1|D)}{p(\mathcal{H}_0|D)}}_{\text{posterior odds}} = \overbrace{\frac{p(D|\mathcal{H}_1)}{p(D|\mathcal{H}_0)}}^{\textbf{Bayes Factor}} \cdot \underbrace{\frac{p(\mathcal{H}_1)}{p(\mathcal{H}_0)}}_{\text{prior odds}}$$

A Bayes factor BF_{10} of 10 means . . .

- The data were 10 times more likely under the alternative model $\mathcal{H}_1$
- We should shift our beliefs by a factor of 10 toward $\mathcal{H}_1$
- If we were evenly split between $\mathcal{H}_0$ and $\mathcal{H}_1$before (prior odds of 1), we have to favor $\mathcal{H}_1$ by 10 after (posterior odds)

Posterior Beliefs about the Hypothesis

$$\mathrm{BF}_{10} = \frac{p(\mathcal{H}_1|D)}{p(\mathcal{H}_0|D)}$$

$$\mathrm{BF}_{01} = \frac{p(\mathcal{H}_0|D)}{p(\mathcal{H}_1|D)}$$

$$\mathrm{BF}_{01} = {}^{1}\!/_{BF_{10}}$$

t-test: Hypothesis

$$\mathcal{H}_0 : \delta = 0$$
$$\mathcal{H}_1 : \delta \neq 0$$

$$\mathcal{H}_+ : \delta > 0$$
$$\mathcal{H}_- : \delta < 0$$

Bayes Factor Transitivity

$$\mathrm{BF}_{0+} = \mathrm{BF}_{01} \cdot \mathrm{BF}_{1+}$$

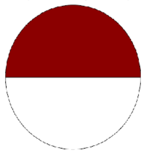
Prove:

$$\frac{p(D|\mathcal{H}_0)}{p(D|\mathcal{H}_+)} = \frac{p(D|\mathcal{H}_0)}{p(D|\mathcal{H}_1)} \cdot \frac{p(D|\mathcal{H}_1)}{p(D|\mathcal{H}_+)}$$

Interpreting Bayes Factor

$p(\text{data} \mid \mathcal{H}_1)$

$\mathrm{BF}_{10} = 1$



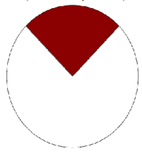
$p(\text{data} \mid \mathcal{H}_0)$

Interpreting Bayes Factor

$p(\text{data} \mid \mathcal{H}_1)$

$\mathrm{BF}_{10} = \frac{1}{3}$

$\mathrm{BF}_{01} = 3$



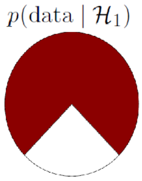
$p(\text{data} \mid \mathcal{H}_0)$

Interpreting Bayes Factor

$p(\text{data} \mid \mathcal{H}_1)$

$\mathrm{BF}_{10} = 3$

$\mathrm{BF}_{01} = \frac{1}{3}$



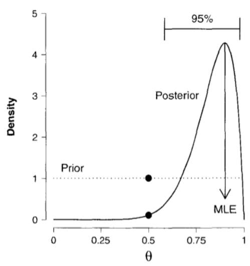
$p(\text{data} \mid \mathcal{H}_0)$

Interpreting Bayes Factor

Evidence categories for bayes factors (Jeffreys, 1961)

Bayes factor BF_{12}			Interpretation
	>	100	Extreme evidence for $\mathcal{M}_1$
30	–	100	Very strong evidence for $\mathcal{M}_1$
10	–	30	Strong evidence for $\mathcal{M}_1$
3	–	10	Moderate evidence for $\mathcal{M}_1$
1	–	3	Anecdotal evidence for $\mathcal{M}_1$
	1		No evidence
1/3	–	1	Anecdotal evidence for $\mathcal{M}_2$
1/10	–	1/3	Moderate evidence for $\mathcal{M}_2$
1/30	–	1/10	Strong evidence for $\mathcal{M}_2$
1/100	–	1/30	Very strong evidence for $\mathcal{M}_2$
	<	1/100	Extreme evidence for $\mathcal{M}_2$

Savage Dickey Method



Example:

- Data: z=9, $N = 10$

Hypothesis:

- $\mathcal{H}_0 : \theta = 0.5$
- $\mathcal{H}_1 : \theta \neq 0.5$

$$BF_{01} = \frac{p(D|\mathcal{H}_0)}{p(D|\mathcal{H}_1)} = \frac{p(\theta = 0.5|D, \mathcal{H}_1)}{p(\theta = 0.5|\mathcal{H}_1)}$$

Priors: (One Sample) t-test

Model

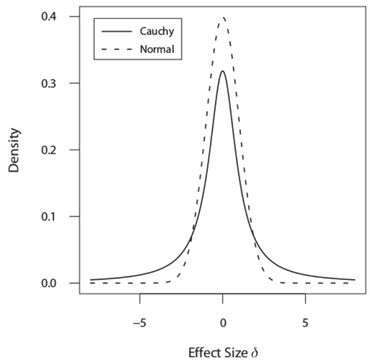
$$y_i \overset{\text{iid.}}{\sim} \text{Normal}(\mu, \sigma^2)$$

$$\delta = \frac{\mu}{\sigma} \quad \rightarrow \quad \mu = \delta\sigma$$

JZS priors

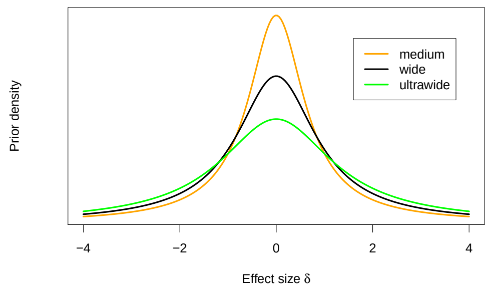
$p(\sigma) \propto \frac{1}{\sigma^2}$	Jeffreys (1961)
$\delta \sim \text{Cauchy}(0, r)$	Zellner & Siow (1980)

Priors: One Sample t-test



Priors: One Sample t-test

BayesFactor package (Morey, 2014) uses different r as prior.
medium: $r = \sqrt{2}/2$ (default!), wide: $r = 1$, ultrawide $r = \sqrt{2}$



JASP-Demo

Two Sample t-test

Model

Group A:

$$x_i \overset{\text{iid.}}{\sim} \text{Normal}\left(\mu + \frac{\beta}{2}, \sigma^2\right)$$

Group B:

$$y_i \overset{\text{iid.}}{\sim} \text{Normal}\left(\mu - \frac{\beta}{2}, \sigma^2\right)$$

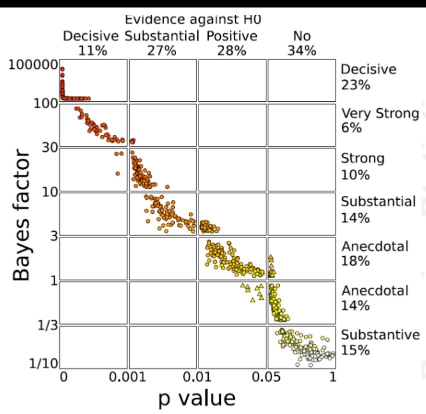
Effect size:

$$\delta = \frac{\beta}{\sigma}$$

Comparing Frequentist & Bayesian t.tests

Wetzels et al., 2011

- 855 t-tests
- in 252 articles



Bayes Factor Package in R

Bayes Factor Package: Tests

Function	Description
ttestBF	Bayes factors for one- and two- sample designs
anovaBF	Bayes factors comparing many ANOVA models (only categorical predictors)
regressionBF	Bayes factors comparing many linear regression models (only continuous predictors)
generalTestBF	Bayes factors for all restrictions on a full general linear model
lmBF	Bayes factors for specific linear models

Bayes Factor Package: Further useful Functions

Function	Description
posterior	Sample from the posterior distribution of the numerator of a Bayes factor object
recompute	Recompute a Bayes factor or MCMC chain, possibly increasing the precision of the estimate
compare	Compare two models; typically used to compare two models in BayesFactor MCMC objects

Exercise: t-test

- Topolinski and Sparenberg (2012): clockwise movements induce psychological states of temporal progression and an orientation toward the future and novelty.
- Concretely: participants who turn kitchen rolls clockwise report more openness to experience.
- Replication data from replication project of Wagenmakers

Thank you