

From Frequentist Problems Towards Bayesian Solutions

Part 2: Introduction to Bayesian statistics (with JASP)

Jorge N. Tendeiro

10 August 2019

Slides at: <https://rebrand.ly/Nagoya2019-Part2>

GitHub: <https://github.com/jorgetendeiro/Nagoya-Workshop-10-Aug-2019>

Today

- Suggested reading material
- Frequentist versus bayes
- Bayes rule
- Jasp
- Summary of todays workshop
- References

Suggested reading material

Introduction to Bayes

Papers:

- Etz & Vandekerckhove (2018): The “Harry Potter” paper.
Very accessible introduction, with examples.
- Etz et al. (2018): “How to become a Bayesian in eight easy steps: An annotated reading list”.
Yes, Alexander Etz writes ‘readable’ papers *very* well. Strongly advised read, but it takes quite some time to process.
- Kruschke (2013): Besides providing an excellent introduction to core concepts, Kruschke offers a discussion over the “testing” versus “estimation” tension. I personally like Kruschke’s position on this matter.

Books:

- Kruschke (2015): The “puppies” book.
Accessible book, with plenty of examples and code. Perfect as a first pick.
- McElreath (2016): From what I read thus far, this book is a jewel.
- Lambert (2018): I’m currently half way. Seems perfect for teaching (hence learning!).
- Gelman (2014): More advanced read (perhaps not the first pick), but truly a master piece.

Learn about JASP

Website

<https://jasp-stats.org/how-to-use-jasp/>.

It includes:

- [Tutorial section](#)
- [YouTube channel](#) to see (and hear) it in action.
- [Forum](#) to post lingering questions.
- [Teaching with JASP](#) includes much more material.

Video tutorials

Etz (who else?) is making videos as we speak.

<https://alexanderetz.com/jasp-tutorials/>.

Papers

- Marsman & Wagenmakers (2017, in *European Journal of Developmental Psychology*).
- Wagenmakers et al. (2018, in a special issue in *Psychonomic Bulletin & Review*).

Frequentist *versus* Bayes

Frequentist paradigm

Concept of **probability**:

- Long-run frequency of a **procedure**.

The probability of a fair coin landing up heads is 50%.

- One cannot state anything about one single event in the long-run sequence.

What is the probability that the **next** coin toss lands heads?

- Recall the definitions of a **p -value** and **confidence interval**: They are based on long-run frequencies.
Conclusion: What can we really conclude from one p -value or one confidence interval?...

Bayesian paradigm

Concept of **probability**:

- Degree of belief.
- Expression of uncertainty about the true state of affairs.
- Subjective: Different people have different beliefs.
- Data are used to update one's belief, by means of the *laws of probability*.
- Applies to both single and repetitive events.

But how do we update our belief in light of data?

Bayes' rule

Bayes' rule

Let \mathcal{A} denote something we want to study.
This can be:

- A parameter, like the mean μ of a population.
- A hypothesis, like $\mu > 100$.

Bayes' rule:

$$p(\mathcal{A}|\text{data}) = \frac{p(\mathcal{A})p(\text{data}|\mathcal{A})}{p(\text{data})}$$

- $p(\mathcal{A})$: **Prior** probability.
- $p(\text{data}|\mathcal{A})$: Data **likelihood**.
- $p(\text{data})$: Marginal likelihood.
- $p(\mathcal{A}|\text{data})$: **Posterior** probability.

Bayes' rule and frequentism

Important:

- The Bayes' rule is a mathematical necessity, it follows from the axioms of probability.
- Frequentists do not dispute this formula!

Bayes' rule and model comparison

Say we have **in total** two competing hypotheses, \mathcal{H}_0 and \mathcal{H}_1 .

We can apply the Bayes' rule to either hypothesis:

$$p(\mathcal{H}_0|\text{data}) = \frac{p(\mathcal{H}_0)p(\text{data}|\mathcal{H}_0)}{p(\text{data})} \quad , \quad p(\mathcal{H}_1|\text{data}) = \frac{p(\mathcal{H}_1)p(\text{data}|\mathcal{H}_1)}{p(\text{data})} .$$

Now divide both equations:

$$\underbrace{\frac{p(\mathcal{H}_0|\text{data})}{p(\mathcal{H}_1|\text{data})}}_{\text{Posterior odds}} = \underbrace{\frac{p(\mathcal{H}_0)}{p(\mathcal{H}_1)}}_{\text{Prior odds}} \times \underbrace{\frac{p(\text{data}|\mathcal{H}_0)}{p(\text{data}|\mathcal{H}_1)}}_{BF_{01}}$$

where BF_{01} is the **Bayes factor**.

The Bayes factor is a measure of the *relative evidence* in the data for either model. E.g., if $BF_{01} = 5$:

- The data are 5 times as likely under \mathcal{H}_0 than under \mathcal{H}_1 .
- After looking at the data, we now support \mathcal{H}_0 five times as much.

Bayes factor

A whole lot can be said about Bayes factors.

They have fervent followers (e.g., Kass & Raftery, 1995; Dienes, 2014; Morey, Romeijn, & Rouder, 2016; E.-J. Wagenmakers et al., 2018).

But there are also critics, including myself (Tendeiro & Kiers, 2019).

JASP

Note of caution

JASP is 'Bayes factor'-oriented.

I personally dislike it, as I think parameter estimation offers a far a clearer, all-inclusive, paradigm.

- To see why I think this, see our preprint: Kiers & Tendeiro (2019).

Worked-out example

Let's jump to JASP now!

The pet example that I will use is the first experiment of Bem (2011):

Precognitive detection of erotic stimuli.

- $n = 100$ (50 men, 50 women), 36 trials per subject.
- In each trial:
 - Two curtains shown side by side.
 - One curtain hides a picture, the other hides a blank wall.
 - Erotic and nonerotic pictures randomly intermixed.

Main research hypothesis:

Subjects are able to “feel” where the erotic pictures are more often than chance (!!!).

Some results from Bem (2011)

Across all 100 sessions, participants correctly identified the future position of the erotic pictures significantly more frequently than the 50% hit rate expected by chance: 53.1%, $t(99) = 2.51$, $p = .01$, $d = 0.25$. In contrast, their hit rate on the nonerotic pictures did not differ significantly from chance: 49.8%, $t(99) = -0.15$, $p = .56$.

The difference between erotic and nonerotic trials was itself significant, $t_{\text{diff}}(99) = 1.85$, $p = .031$, $d = 0.19$.

(...) the correlation between stimulus seeking and psi performance was .18 ($p = .035$).

Descriptive classification of Bayes factors

BF_{10}	Qualitative descriptive
1	No evidence
1 — 3	Anecdotal evidence for \mathcal{H}_1
3 — 10	Moderate evidence for \mathcal{H}_1
10 — 30	Strong evidence for \mathcal{H}_1
30 — 100	Very strong evidence for \mathcal{H}_1
> 100	Extreme evidence for \mathcal{H}_1

Source: Lee & Wagenmakers (2013)

Note: Do not take these qualitative labels strictly. Use them as loose reference bounds.

Summary of today's workshop

Main ideas: Replication crisis

Plenty of problems in psychological research have been identified throughout the years:

- Results do not replicate.
- Bias.
- QRPs.
- CIs and p -values poorly understood and often misused.

Psychology is currently in the middle of a revolution. Several solutions are being worked out:

- Preregistration.
- Registered reports.
- Open data, materials.
- Embrace uncertainty. Avoid dichotomous thinking.
- Better statistical analyses. In particular: **Stop using NHST.**

The entire research ecosystem is picking up on these changes **fast!**

Main ideas: Bayesian statistics

Bayesian statistics is gaining traction as a viable alternative.

- Quantification and accumulation of evidence.
- Logical updating of belief.
- Avoid long-standing fallacies of classical statistics.

JASP in particular is one very friendly software that can ease the use of Bayesian statistics.

JASP is mostly model comparison/ hypothesis testing based.

I suggested that there is valid criticism against *only* hypothesis testing (e.g., Tendeiro & Kiers, 2019; Kiers & Tendeiro, 2019).

Beyond JASP, I advocate model fitting, parameter estimation, and reporting summaries of posterior distributions.

Tools needed: MCMC sampling (e.g., JAGS, Stan).

References

- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100(3), 407–425. doi: [10.1037/a0021524](https://doi.org/10.1037/a0021524)
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 781. doi: [10.3389/fpsyg.2014.00781](https://doi.org/10.3389/fpsyg.2014.00781)
- 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. doi: [10.3758/s13423-017-1317-5](https://doi.org/10.3758/s13423-017-1317-5)
- Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian Inference for Psychology. *Psychonomic Bulletin & Review*, 25(1), 5–34. doi: [10.3758/s13423-017-1262-3](https://doi.org/10.3758/s13423-017-1262-3)
- Gelman, A. (2014). *Bayesian data analysis* (Third edition). Boca Raton: CRC Press.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795. doi: [10.2307/2291091](https://doi.org/10.2307/2291091)
- Kiers, H., & Tendeiro, J. (2019). *With Bayesian Estimation One Can Get All That Bayes Factors Offer, and More* [Preprint]. doi: [10.31234/osf.io/zbpmv](https://doi.org/10.31234/osf.io/zbpmv)
- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2), 573–603. doi: [10.1037/a0029146](https://doi.org/10.1037/a0029146)
- Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (Edition 2). Boston: Academic Press.
- Lambert, B. (2018). *A student's guide to Bayesian statistics*. Los Angeles: SAGE.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge ; New York: Cambridge University Press.
- Marsman, M., & Wagenmakers, E.-J. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology*, 14(5), 545–555. doi: [10.1080/17405629.2016.1259614](https://doi.org/10.1080/17405629.2016.1259614)
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton: CRC Press/Taylor & Francis Group.
- Morey, R. D., Romeijn, J.-W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, 72, 6–18. doi: [10.1016/j.jmp.2015.11.001](https://doi.org/10.1016/j.jmp.2015.11.001)
- Tendeiro, J. N., & Kiers, H. A. L. (2019). A review of issues about null hypothesis Bayesian testing. *Psychological Methods*. doi: [10.1037/met0000221](https://doi.org/10.1037/met0000221)
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25(1), 58–76. doi: [10.3758/s13423-017-1323-7](https://doi.org/10.3758/s13423-017-1323-7)
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Gronau, Q. F. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25, 35–57. doi: [10.3758/s13423-017-1343-3](https://doi.org/10.3758/s13423-017-1343-3)