UNIVERSITÄT BIELEFELD
Fakultät für Wirtschaftswissenschaften

Moving to a World Beyond $p < 0.05$

Reading Course, Winter Term 2020/2021

Key points and questions

Agenda

1. Key points (as collected by you)
2. Key points (added by me)
3. Questions (as collected by you)
4. Miscellaneous

Specific advice for data analysis...

- Focus on a thorough description of data and methods.
- It's important to describe the chosen approach in detail and explain all your decisions and why/how you made them and how you finally got to the results and interpretation.
- Don't get hung up on a threshold.
- Don't default to any method just because it is common practice or conveniently implemented.
- There are many alternatives to the traditional approach.

... and interpretation of results

- The context and background should always be kept in mind. (Does the interpretation make sense?)
- Think about how exactly the method works and under which assumptions, how its results can be interpreted, and why that fits your needs and the data you want to apply it to.
- It is not about the quality of methods but about the quality of their interpretation.
- View any single study in the context of all previous studies done on the topic. Studies don't necessarily disprove each other, they all add to a growing knowledgebase.
- Do not conclude anything from a single study with unwarranted certainty.

Open science

- Be open and describe not only all analyses performed but also present all findings regardless of statistical significance.
- Science should be reproducible and scientists should be open to Open Science.

More general remarks

- “Accept uncertainty. Be thoughtful, open, and modest.” $=>$“ATOM.”
- Statisticians cannot address the problem alone. Any approach involving only statisticians will not succeed. Scientists from other fields are needed for reform movements.

Some complementary points from my side

- Use and publish pre-analysis plans (if appropriate).
- Define and discuss (beforehand if possible) meaningful effect sizes.
- Clarify your statistical goals explicitly and unambiguously.
- Define precisely the population of interest in research studies and carefully assess whether the data being analyzed are representative of the population (wrt to the variable you are interested in).
- Clearly differentiate between confirmatory and exploratory research.

Some complementary points from my side

- Report descriptive statistics and visualize your data (and results).
- Ban asterisks indicating statistical significance.
- Avoid reporting p-values with inequalities (e.g. $p < 0.05$).
- Stop using the term “statistically significant” entirely.
- Consider (reverse) Bayesian approaches.
- Check the assumptions of your models.
- Look for and present results from many models that fit the data well.

Some complementary points from my side

- Tolerate and communicate uncertainty in all statistical conclusions, seeking ways to quantify, visualize, and interpret the potential for error.
- Evaluate the strength of empirical evidence based on the precision of the estimates and the plausibility of the modeling choices.
- Consider the scientific context, related prior evidence / expert judgement, plausibility of mechanism, study design and data quality, real world costs and benefits, novelty of finding, and other factors.

Some complementary points from my side

- Accept that one study is rarely definitive, so encourage, sponsor, conduct, and publish replication studies.
- Use meta-analysis, evidence reviews, and Bayesian methods to synthesize evidence across studies.
- **Read Wasserstein et al. (2019).**
- **Read van Dongen et al. (2019) for getting some good examples of statistical analyses.**
- **Bear in mind: There is no one fits all solution!**

Questions

- Wasserstein et al. write “If you use statistics in research, business, or policymaking but are not a statistician, these articles were written with YOU in mind” - I don’t think that is true. To me, this feels much more like a discussion among (academic) statisticians and I don’t think we can possibly hope that practitioners will change their behavior until there is as much agreement about the do’s as there is about the don’ts of statistic analyses in the statistical community. Or am I to pessimistic, underestimating the progress of the discussion?
- We have seen so many different solutions to the problem of interpreting the p-values, and none of them has really been able to prevail so far. Now I ask myself why. The practical implementation is completely missing so far, isn’t it? I think most of the articles are from 2018. why isn’t there already progress in implementing a new concept in practice, that improves the interpretation of p-values?
- Can we expect to see a shift away from p-values or is the scientific community too reliant on it/ used to it at the moment?

Questions

- There are so many different statistical methods and it seems like none of them is the best way to go. Should journals aim to publish as many different study approaches as possible and examine one topic from different point of views instead of focusing on significance and different topics?

Questions

- How can one convince journal editors to change their practices?
- In the context of “publish-or-perish”, are there potential solutions? How to change today’s practice of producing an endless number of papers with little to zero contribution to world knowledge?
- How should one design/change statistics education specifically/Are there widely accepted guidelines? Have changes already been made?

Questions

- Is there an overview/systematic analysis of institutional causes for the current problems?
- Is there an overview of journals with nonsensical requirements for submissions?

Questions

- Are we part of the problem and if yes, how can we become part of the solution?

Miscellaneous

Schedule

Date	Paper
02.11.20	Organisation & Introduction
	Interpreting and using p
09.11.20	Betensky – The p-Value Requires Context, Not a Threshold
16.11.20	Greenland – Valid P-Values Behave Exactly as They Should: Some Misleading Criticisms of P-Values and Their Resolution With S-Values
	Supplementing or replacing p
23.11.20	Goodman et al. – A Proposed Hybrid Effect Size Plus p-Value Criterion: Empirical Evidence Supporting its Use
30.11.20	Benjamin et al. – Three Recommendations for Improving the Use of p-Values
07.12.20	Colquhoun – The False Positive Risk: A Proposal Concerning What to Do About p-Values
14.12.20	Matthews – Moving Towards the Post $p < 0.05$ Era via the Analysis of Credibility
	Adopting more holistic approaches
21.12.20	Billheimer – Predictive Inference and Scientific Reproducibility
04.01.21	Amrhein et al. – Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication
11.01.21	McShane et al. – Abandon Statistical Significance
18.01.21	Ziliak – How Large Are Your G-Values? Try Gosset's Guinnessometrics When a Little "p" Is Not Enough
25.01.21	van Dongen et. al – Multiple Perspectives on Inference for Two Simple Statistical Scenarios
	Optional
01.02.21	Wasserstein et al. – Moving to a World Beyond "$p < 0.05$"
08.02.21	Biontech Study (voluntary)

Schedule

- Grades will be transferred to the examination office this week.
- Who is likely to participate next week?
- Who volunteers to give a short presentation next week?

Presentations

- It may be cool to upload the presentations (and associated R Code) to github such that they are publicly available.
- If and only of you agree to upload your presentation, please drop me a chat message or a mail. You may of course delete your names in the presentations, I wouldn't upload the material before next week.

Outlook

- If your want to write your master's thesis about meta-research, statistical inference, journal policies, want to conduct a replication study or the like, you can approach me.

Looking Back: Initial aims of this course

- At the end of the course, you should...
  - ... be critical regarding common practices in empirical research (e.g., using holy p-value thresholds to summarize results of a study)
  - ... know about alternatives in statistical inference
  - ... have an insight into a currently discussed hot topic in empirical research
  - ... have some experience how to present research papers
  - ... still love statistics ;)
- **Any feedback is welcome!**

Thanks a lot! It was a pleasure to work with you!