

STAT 518/539 Designing the experiments and presenting the results

Marina Meilă

Department of Statistics
University of Washington

April 3, 2015

Why do/present experiments at all?

The pragmatic view

- most journals/conferences require them (even when not necessary!)

Why do/present experiments at all?

The pragmatic view

- most journals/conferences require them (even when not necessary!)

The naive view

- Look, it actually works!

Why do/present experiments at all?

The pragmatic view

- most journals/conferences require them (even when not necessary!)

The naive view

- Look, it actually works!
- Look, it actually works on data **THIS BIG!**

Know what is the point you want to make

- Every experiment should answer a question (or more)
- Every graph presented in a talk/paper should answer a question (or more)

Logical flow

- Questions
- Experiments
 - What experiments will best answer the questions?
 - What questions can be answered by experiment?
 - Time, report length management: what experiments are most important? practical?
- Graphs
 - What graph best displays the answers/features I want to show?
- Discussion
 - Interpret the graphs to spell out the answers to the original questions
 - Draw attention to surprising features of the results
 - Does the graphs/experiments raise new questions?
 - To what extent is the experiment conclusive?

Questions: Algorithm analysis

- Do the theoretical predictions hold?
 - “Feel good” experiments, but less interesting
 - ... unless the predictions depend on unknown constants
 - ... unless the assumptions of the theory cannot be verified in practice (e.g. is the method robust to violation of the assumptions)
- Testing for robustness
 - Noise and other violations in known amount: when does the algorithm break? slow down?
 - Which assumptions are critical and which aren't
- What component of the algorithm is essential to the performance/accuracy?
 - e.g. test with/without a computational “improvement”
- How does the performance scale? (what influences the running time/memory)

Questions: Empirical comparisons

Is my algorithm “better”?

- Better must be made precise: more accurate, faster, requires less data, more robust, easier to implement and understand
- Which algorithm is better may depend on the current data. Test in a variety of experimental conditions

Figures and experimental results

- Show only relevant results
- General principles for displaying information (figures vs. tables, etc), see Tufte
- Make clear annotations, as close to data as possible. Spare comments (e.g. “lower is better”)

Rule Figure/Table caption explains **what** is in the figure/table; comments are in the main text. Captions must be self-contained.

- Discussing the experimental results
 - don't repeat what's in the tables/figures (but refer to them)
 - appeal to reader's intuition
 - but be precise