Homework 7. Due by 11:59pm on Sunday 10/26.

Negligence, mistakes & how to avoid them

Researchers, like all other humans, make mistakes. Read pages 12–14 of *On Being a Scientist.* Write brief answers to the following questions, by editing the tex file available at https: //github.com/ionides/810f25, and submit the resulting pdf file via Canvas.

1. How should one balance the professional consequences of errors with the professional requirement to publish?

   It is impossible for researchers to completely avoid errors. To fulfill their requirement to publish, they should stick to scientific disciplines. Every scientific result must be carefully prepared, submitted to the peer review process, and scrutinized even after publication(quoted from On Being A Scientist).

2. What is a reasonable level of skepticism about correctness of published results?

   There should not be too skeptical, because published results are often strictly filtered, reviewed, and corrected immediately when there are typos. Such trust can significantly improve efficiency. However, when they find errors during implementing and utilizing results from these work, it is reasonable to doubt them.

3. If you think you have identified an error in someone else's published paper, what are possible courses of action? What are their advantages and disadvantages?

   When such a mistake appears in a journal article or book, it should be corrected in a note, erratum (for a production error), or corrigendum (for an author' error). Mistakes in other documents that are part of the scientific record-including research proposals, laboratory records, progress reports, abstracts, theses, and internal reports-should be corrected in a way that maintains the integrity of the original record and at the same time keeps other researchers from building on the erroneous results reported in the original(quoted from On Being A Scientist).

   Advantages: Formality, integrity, procedural justice(mostly positive).

   Disadvantages: It might have a negative influence on researchers' reputation.

4. Look on the internet in a leading statistics journal (such as Journal of the American Statistical Association and Annals of Statistics) for papers with a corrigendum, erratum or retraction. How common is it in our discipline to publish corrections of mistakes? If you find it is rare, how does the field avoid building on incorrect results?

   Yes, it is rare. It is so because every scientific result must be carefully prepared, submitted to the peer review process, and scrutinized even after publication.

5. What kinds of errors arise in statistics research, and what are good research practices to avoid or reduce them?

(a) in theoretical results;

Derivation errors; Example Errors;

To avoid them, one needs to be careful in every step in their derivation, and choose examples carefully.

(b) in numerical results.

Data does not represent population; Assumption Errors:

To avoid them, one should try to choose data that includes both normal cases and extreme cases. They should also make reasonable assumptions.

6. **A capstone question**. This is the last homework on the topic of responsible conduct in research and scholarship. Here is a final RCRS question, which concerns all the classes 1–8. We have now discussed various incentives for collegial behavior and consequences for antisocial behavior, in the context of research, teaching and professional service in our field. Can you think of situations where the consequences for RCRS violations are:

(a) Too lenient. Consequences insufficient to effectively disincentivize antisocial behavior; or too much burden of evidence required to impose consequences; or strong incentives not to impose adequate consequences.

I think the consequences for conflict of interests are too lenient. We can still see researchers who are interested in industry let their mentees do projects relevant to their own ambition.

(b) Too harsh. Disproportionately damaging consequences for minor violations; or penalties imposed with too little burden of evidence; or no presumption of innocence; or incentive structures that lure people into breaking rules and then punish them for it.

I think the requirement for public data is too harsh. In my humble view, I still think publishing data after the research is done shall be fine. Sometimes, publishing data before the research is done might potentially provide advantages to competitors who are working on similar ideas, which, I think, might be unfair for researchers.