Reproducible Research
The literate programming solution

Susan Holmes
http://www-stat.stanford.edu/~susan/
@SherlockpHolmes

Statistics Department, Stanford University
Fellow Center for the Advanced Study of the Behavioral Sciences

MSRI, February-March, Complementary program, 2018

Mathematical versus Statistical Scholarship

$(A \rightarrow B) \iff (\neg B \rightarrow \neg A)$

Observation: Non B ($\neg$B)
Conclusion: Observing $\neg$B,
gives A is not true.

Statistical Logic:

$(A \rightarrow B) \iff (\neg B \rightarrow \neg A)$

If the observed X makes
$P(B)$ very small, then we
infer A is unlikely.

In mathematics: Having *a proof* means to has been verified by at least (1/2/3) Phd mathematicians in the field.

In computational science: *"An article in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures."* Jon Claerbout,1992

In computational statistics: all *the instructions, the data and the explanation of the methods* used are collected so that any reader can reproduce the study.

The R and Bioconductor projects allow this.
Literate Programming, Don Knuth, 19992.
See Statistical Proof? in January 2018, Bulletin of the AMS.