Principles of Scientific Research

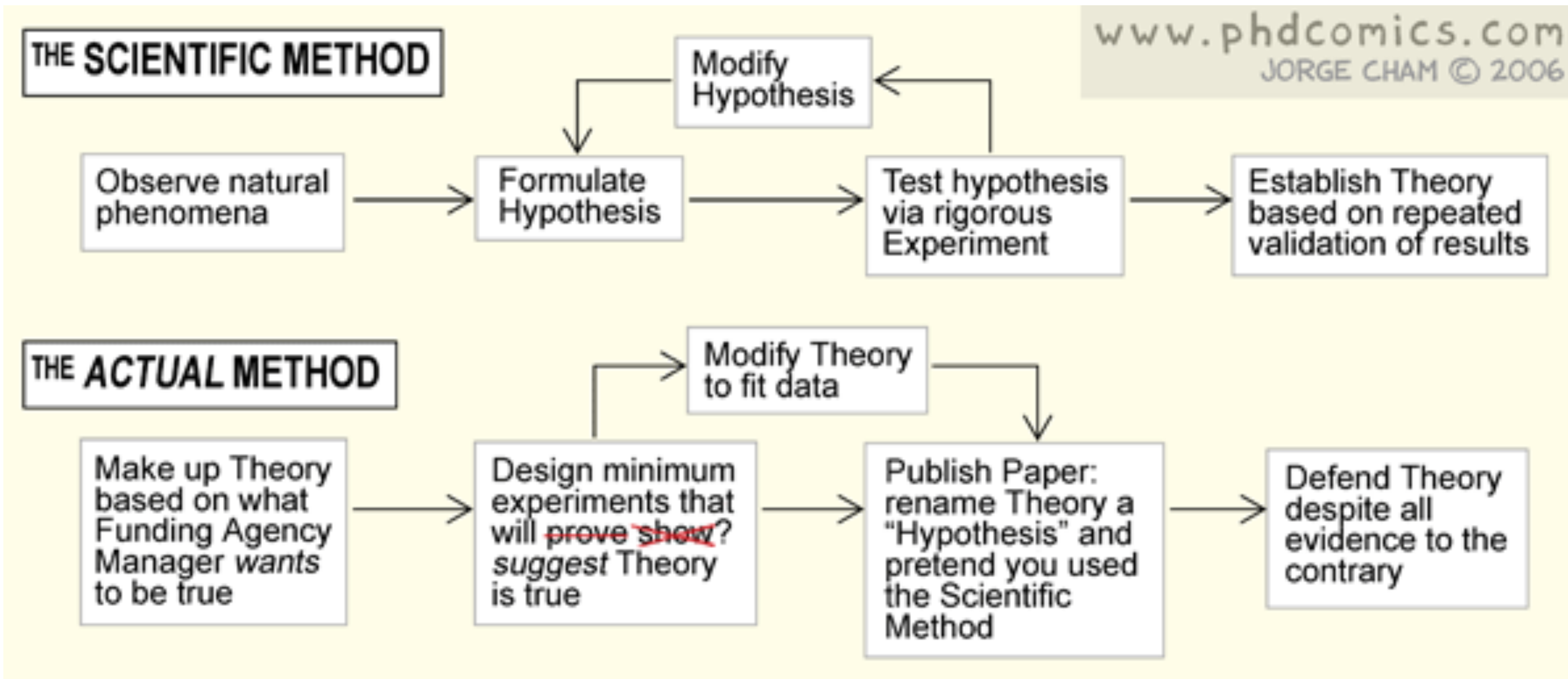
Formulating and testing hypotheses

Resources

Chapters 1 & 2, Valiela
(Chapters 3 & 4 contain excellent discussions of applying statistics and more detail on study design)

Chapters 5 & 7, Smith

"Designing a Research Project",
Ch. 6 of Introduction to Educational Research
http://wps.ablongman.com/ab_mertler_edresearch_6/)


www.phdcomics.com
JORGE CHAM © 2006
THE SCIENTIFIC METHOD
Modify Hypothesis
Observe natural phenomena
Formulate Hypothesis
Test hypothesis via rigorous Experiment
Establish Theory based on repeated validation of results
THE ACTUAL METHOD
Modify Theory to fit data
Make up Theory based on what Funding Agency Manager wants to be true
Design minimum experiments that will ~~prove~~ ~~show~~? suggest Theory is true
Publish Paper: rename Theory a "Hypothesis" and pretend you used the Scientific Method
Defend Theory despite all evidence to the contrary

Deductive vs. empirical science

- Galileo:
  empirical observations are used to obtain factual information, then deductive work synthesizes observations into laws
- Deductive science
  - Use of constructs / relationships to logically relate things
    - "Tautologies" are models useful to explain and integrate observations
  - Typical progression:
    - Model ➔ predictions ➔ experiment
      - Do predictions and experiments agree?
    - Roles of models
      - Make predictions to test
      - Help evaluate whether knowledge of something is sufficiently complete
      - Numerical experiments save resources, time (sometimes are only possible way to test)
- Empirical science
  - Basic principles
    - Testability (empirical verification)
    - Operational definition (makes ideas testable)
    - Controlled observations

Methods of conducting scientific research

- Consider each of these in turn:
  - Observation
  - Hypothesis
  - Experimentation
  - Interpretation

- Observation
  - Purposeful
  - Relate to other things you know, but don't exclude if doesn't (yet) fit
  - Be equally critical of "good" and "bad"
  - Tools are vital for observations, but recognize potential errors / limitations
- Hypothesis
  - "An imaginative preconception of a factual relationship" (Smith)
  - Theoretical generalizations, vs. empirical generalizations that are summary statements of fact
  - Must be testable, in fact, generally try to disprove an hypothesis
  - *More on hypotheses later….*

Principles of Scientific Research (continued)

- Experimentation
  - Especially for hypothesis testing: don't just "try it"
    - "pilot work" is good, but generally less rigorous and data are "preliminary" – make sure the time and effort go into definitive experimentation
  - Valiela discusses utility of manipulative experiments
    - Linchpin of the "classical scientific method"
    - Test effects of an independent variable on dependent variable
    - Need control experiments!
  - But sometimes have to rely on correlational studies
    - Correlations can only suggest variable vary together, NOT a causal link
    - Comparative studies are similar idea, but find different conditions for independent variable
    - Conclusions most valid for data set *on aggregate*
  - Written protocols, lab notebooks document work
- Interpretation
  - What do the data "mean"?
  - Statistics
  - Graphing / tabulation

More on Hypotheses

- Hypotheses must be
  - Specific and testable
  - Based on the known, extending to unknown
- Null hypothesis: no cause and effect, related only by chance
  - Often easier to test than original hypothesis
- Hypotheses help to structure experimentation and data collection
- Methods to develop hypotheses:
  - Method of agreement
    - If an event is repeated, with one factor in common, then factor may be cause of event
  - Method of difference
    - If an event is repeated with one factor but not another, first factor is causative agent
  - Concomitant variation
    - In an increase in the intensity of a factor is followed by parallel variation in event, factor is cause of event
  - Note none of these are definitive and require controlled experiments to test
- Experimental design and statistics critical to hypothesis testing
  - Statistics lab on campus is important resource! NO CHARGE for consulting
  - http://www.stat.colostate.edu/consulting/stat_lab.html
- Remember Occam's razor

Type I and II errors

- In tests of hypotheses, suppose you have 2 populations and you want to determine if they are statistically different. Null hypothesis = no, they are the same.
  - Any test of the hypothesis yields a continuous *range* of probabilities (highly likely to highly unlikely)
  - We generally use "significance", p=0.05 or p=0.01, to decide on test outcome (let α be equivalent to "p")
- Our hypothesis could be true or false, and our test can accept or reject it
  - Suppose it was in fact true, but our test rejected it?
    → Type I error
    - Avoid by demanding less uncertainty? That increases the Type II error rate
  - Suppose it was in fact false, but our test accepted it?
    → Type II error
    - The probability of this kind of error is β and increases as two means come closer to each other
    - "power" is 1-β and refers to probability of rejecting the null hypothesis when it is indeed false
    - Power can be increased by further data collection

Hypotheses are inherently un-provable

- Valiela makes the point that only "operational" concepts can be tested
  - Inability to test existence of "ether"
- Can *disprove* hypotheses
  - Can in fact eliminate a series of hypotheses
  - Those that we fail to disprove are incorporated into the body of knowledge
  - Curry at al.: "Science can help us approach the truth progressively, but we can never be certain that we have arrived at the final explanation."
- So frequently use null hypothesis testing to make progress

Curry et al., 2005: Hypothesis formulation and testing

We therefore have the following several objectives in writing this paper:

- to clarify the debate surrounding the subject as to whether or not global warming is causing an increase in global hurricane intensity, by sorting out the valid from the fallacious criticisms, addressing the valid criticisms, assessing alternative hypotheses, and identifying the outstanding uncertainties;
- to illustrate a methodology of hypothesis testing to address multiple criticisms of a complex hypothesis that involves a causal chain; and
- to provide a case study of the impact of politics, the media, and the World Wide Web on the scientific process.

This paper is valuable for showing how effective rebuttals can be made against both valid and invalid criticisms, and how proper hypothesis formulation and testing is needed for results to stand up to peer review

- Central hypothesis: greenhouse warming is causing an increase in global hurricane intensity
  - Can't be tested because the multi-decadal, global dataset needed does not exist, AND global climate models cannot (yet) resolve the critical hurricane processes
  - Approach: formulate as a causal chain of three SUB-hypotheses that CAN be tested
- In rebuttal, divided issues into
  - Those having obvious logical fallacies
  - Those having logical fallacies, but raising ancillary issues that need to be addressed
  - Concerns based on logically valid argument

LOGICAL FALLACIES

The logical fallacies discussed in the context of attacks on the hypotheses are summarized here, following elementary texts in logic, such as Damer (2004) [the Wikipedia is an excellent online source describing these fallacies (online at http://en.wikipedia.org/wiki/Logical_fallacy)].

- An *ad hominem fallacy* consists of asserting that someone's argument is wrong because of something discreditable/not authoritative about the person or persons cited by them rather than addressing the soundness of the argument itself.
- *Appeal to authority* cites a person or organization who is an authority in the relevant field and therefore should carry more weight, but given the possibility of mistake, should not be compelling. Direct evidence must be provided by the expert and the expert should be reasonably unbiased.
- *Appeal to motive* is a pattern of argument that consists of challenging a thesis by calling into question the motives of its proposer.
- An *unrepresentative sample* is one that is falsely taken to be typical of a population from which it is drawn.
- *Begging the question* is a fallacy occurring in deductive reasoning in which the proposition to be proved is assumed implicitly or explicitly in one of the premises.
- *Correlation implies causation* is a logical fallacy by which two events that occur together are claimed to be cause and effect.
- A *fallacy of distribution* occurs when an argument assumes that what is true of the members is true of the class (composition), or what is true of the class is true of its members (division).
- *Hasty generalization* is the logical fallacy of reaching an inductive generalization based on too little evidence.
- *Statistical special pleading* occurs when the interpretation of the relevant statistic is "massaged" by looking for ways to reclassify or requantify data from one portion of results, but not applying the same scrutiny to other categories.
- *Fallacy of the single cause* occurs when it is assumed that there is one simple cause of an outcome when in reality it may have been caused by a number of only jointly sufficient causes.

Example null hypothesis statement and testing

- "There has been no global increase in hurricane intensity over the period from 1970-2004."
  - Argument from detractors: "Category 3 hurricanes cannot be distinguished from Cat 4 and 5, therefore null hypothesis cannot be rejected"
    - Curry et al.: present arguments about level of uncertainty in classifications → not high enough to confuse Cat 1&2 with Cat 4&5, so must reject null hypothesis
  - But they do note a clear need for new and better datasets, contributed by those most knowledgeable of the field and the observations

Summary

- "If the central hypothesis is to be elevated to a theory, it must pass the following three tests:
  1. Survive scrutiny and debate, including attacks by skeptics,
  2. Be the best existing explanation (physical and statistical) for the particular phenomenon, and
  3. Demonstrate predictive capability."
- "…the central hypothesis and sub-hypotheses cannot be invalidated by the available evidence. We anticipate that it may take a decade for the observations to clarify the situation as to whether the hypothesis has predictive ability. In short, time will tell."

De Laat, BAMS, 2007:

"Recently, Curry et al. (2006) published an interesting paper on the issue of whether greenhouse warming is causing an increase in hurricane intensity. The authors identified 14 distinct critiques on two recent papers (Emanuel 2005; Webster et al. 2005), and investigated whether any logical fallacies were involved. **The use of elementary logic to dissect lines of reasoning and argumentation is a powerful tool, as evidenced by this paper, and the authors should be applauded for using them in such a distinctive manner.** This methodology is more popularly known as "critical thinking," and Curry et al. (2006) mention a number of logical fallacies, although many more exist (Caroll 2000; Haskins 2006).

However, it should be noted that **the authors have an important logical fallacy themselves. …**

The line of reasoning here is that natural factors alone cannot explain the observed twentieth-century temperature variations, while including greenhouse gases does. The logical fallacy is the "fallacy of false dilemma/either-or fallacy," that is, the number of alternatives are (un)intentionally restricted, thereby omitting relevant alternatives from consideration (Haskins 2006)."

Curry, 2011: Nullifying the climate null hypothesis

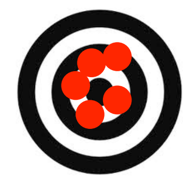
"This essay addresses Trenberth's statement that 'Given that global warming is "unequivocal", and is "very likely" due to human activities to quote the 2007 IPCC report, the null hypothesis should now be reversed, thereby placing the burden of proof on showing that there is no human influence.' **We examine how the concept of a null hypothesis is being used implicitly and explicitly in the scientific and policy debate on climate change, in the context of scientific hypothesis testing**, as a framework for 'burden of proof' arguments and policy deliberations, and metaphorically in the context of a polemic. It is argued that the statement of a null hypothesis is not particularly useful in the broader context of the scientific inferences surrounding the topic of the attribution of climate change and also policy decisions."

Other useful stuff in Valiela Ch. 2

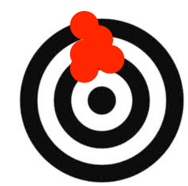
- Frequency distributions
  - Important parameters
  - Assumption of the normal distribution
    - Data transformations can be used to make this assumption true
  - The Poisson distribution
- Data filtering
    - Running means
    - Time series analyses

The value of data collected

- Uncertainty
  - Range within which the value expressed is expected to lie.
- Error
  - The difference between the expressed value and the true value.
- Accuracy
  - The closeness of the expressed value to the true value.



- Precision
  - The closeness of a series of replicate expressions to each other.



- Repeatability
  - Results from the same operator, using the same equipment, etc.
- Reproducibility
  - Different operator(s), different equipment, etc.

Source: Quality in the Analytical Chemistry Laboratory, Elizabeth Prichard, 1995

Moving on to the practical:

- Getting oneself organized to plan and conduct original research

Tasks to be accomplished when planning research

1. State the title, problem, and hypotheses
   - Make sure problem can be investigated scientifically (is well formulated)
   - State the purpose, why topic is important enough to be researched
   - State hypothesis (or null hypothesis)
2. Identify keywords; outline the library search for related information
3. Identify data needed and sources
4. List steps to be carried out to complete the study
5. Specify procedures and tools to be used in collecting data
   - How will you organize data?
   - How will you ensure data quality?
6. Foresee how data can best be analyzed and interpreted
   - Consult Statistics Lab? Other useful consultations before beginning?
7. Anticipate report format most appropriate for your research
   - For us, journal whose mission is consistent with work to be published (e.g., see http://www.ametsoc.org/pubs/journals/)

Literature Surveys (Ch. 7, Smith)

- Reading helps researchers by
  - Stimulating ideas
  - Improving and organizing knowledge
  - Avoiding duplication of previously accomplished work
  - Reinforcing or refuting hypotheses
- Other ways to stay on top of developments in your field:
  - Sign up for automated alerts about new papers
    - http://www.atmospheric-chemistry-and-physics.net/alerts_and_rss_feeds.html
  - Journal clubs
  - Attending seminars, conferences, workshops

Web of Science and Science Citation Index

- Accessing the database from campus and from home
- Keyword, author searches
- Looking up papers that were referenced by the article you are interested in
- Finding papers that referenced the article you are interested in
- Alerts Service:
  - –On the Web of Science main page, fill out the Search parameters as you wish and click on Search
  - –On the results page click on Search History (A sub tab of the Web of Science Tab)
  - –Choose the search that you want an alert for, and click on Save History/Create Alert
  - –Fill out the form that comes up (you can have the current results emailed to yourself too)
  - –Click Done
  - –You can see/modify the results of your alert if you click on Open Saved History

More on searching (from Arsineh)

WOS Citation Alerts –If you would like to see who cites a particular paper, search for it, click on it and then click on Create citation alert on the right hand corner of the abstract (under Times Cited).

Google Scholar searches

Google Reader: You can subscribe to the RSS feed of different journals and they have the option of limiting what shows up for some journals, using keywords.