Designing Experiments

Administrativia

- Be careful about attribution on slides
 - (I should be better, too, but probably won't)
- NSF GRFP
- Role of the subject area advisor
 - Definitely okay to coauthor a paper!
 - Research project should be *developed* by you
 - Doesn't have to be solely your idea
- Starting week 6 (see syllabus)



By Robert Boyle [Public domain], via Wikimedia Commons

Terminology [1]

- Factors, *i.e.*, independent variables or knobs
 - Levels can be qualitative or quantitative
- Response, *i.e.*, dependent variables
- Full vs. fractional factorial
 - Treatments or runs
- Bias or distortion
 - Randomization vs. blocking
- Replication vs. repetition

Example

- My hypothesis: international grad students can drink a cup of hot tea faster than domestic grad students can.
- 1 factor, 2 levels per factor, full factorial
- Experiments run from 11am to 4pm, outside in early August
 - Randomization?
- Timed ten each (international *vs.* domestic), now what?

Student's t test [2] (no pun intended)

- William Sealy Gosset, pen name "Student"
- Worked for Guinness Brewery
- A way to compare 2 Gaussian distributions
- You must reject the null hypothesis in order to be conclusive about anything
 - My null hypothesis: there is no difference between how fast international vs. domestic grad students can drink hot tea

You can look up the fancy math online...



Just an example

- You're unlikely to be asking a simple enough research question to be able to use a t test
- The point is, you should have some idea about what you're going to do and how...
 - Where will you get the data?
 - What will you do with it?
 - What kinds of bias, distortion, and other problems should you be careful about?

Assignment #4

- Due on Friday, September 28th
- Read "The Truth Wears Off", which I'll send out by email
- Be ready to discuss it on Friday
 - (No need to turn anything in)

Ethics (preview?)

- IRB
- Bellmont report
 - Personal data or intervention

Assignment #5

- Due Monday, October 1st
- Write as much as you need to (½ a page to 2 pages is probably good, but that's just a rule of thumb)
- Tell me *exactly* what you plan to do for your semester project in as much detail as possible
 - Where will you get the data? What code will you use? What code will you write? What's your goal? How long will it take? What bias/distortion or other problems will you need to account for? What will you do with the output? What will you archive (and how) to make it as reproducible as possible?

References

 [1] http://www.cs.unm.edu/~crandall/netsfall15/lectur enotes/doeintrochapter1.pdf

An Introduction to DOE by Barrentine, see https://g.co/kgs/rJkpKW

• [2]

https://en.wikipedia.org/wiki/Student%27s_t-test