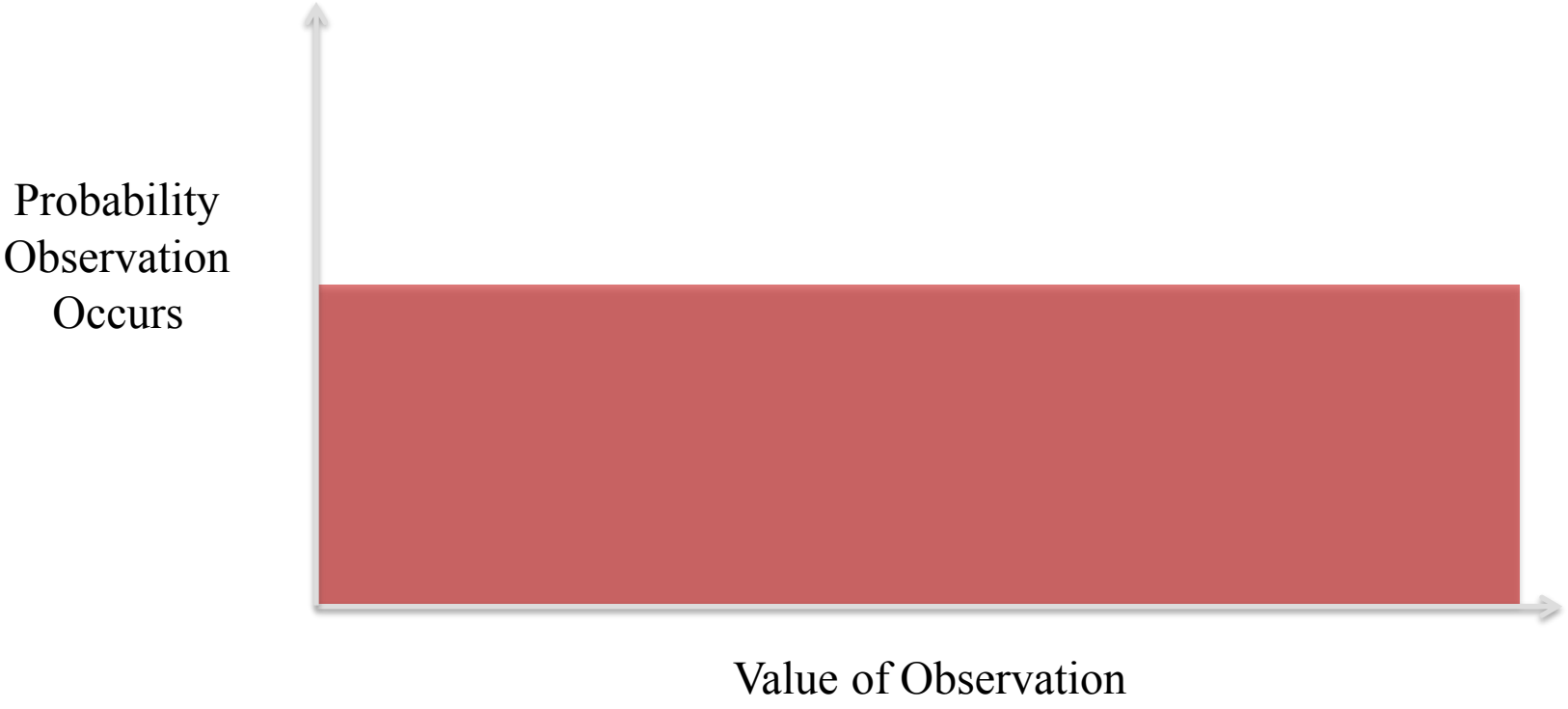
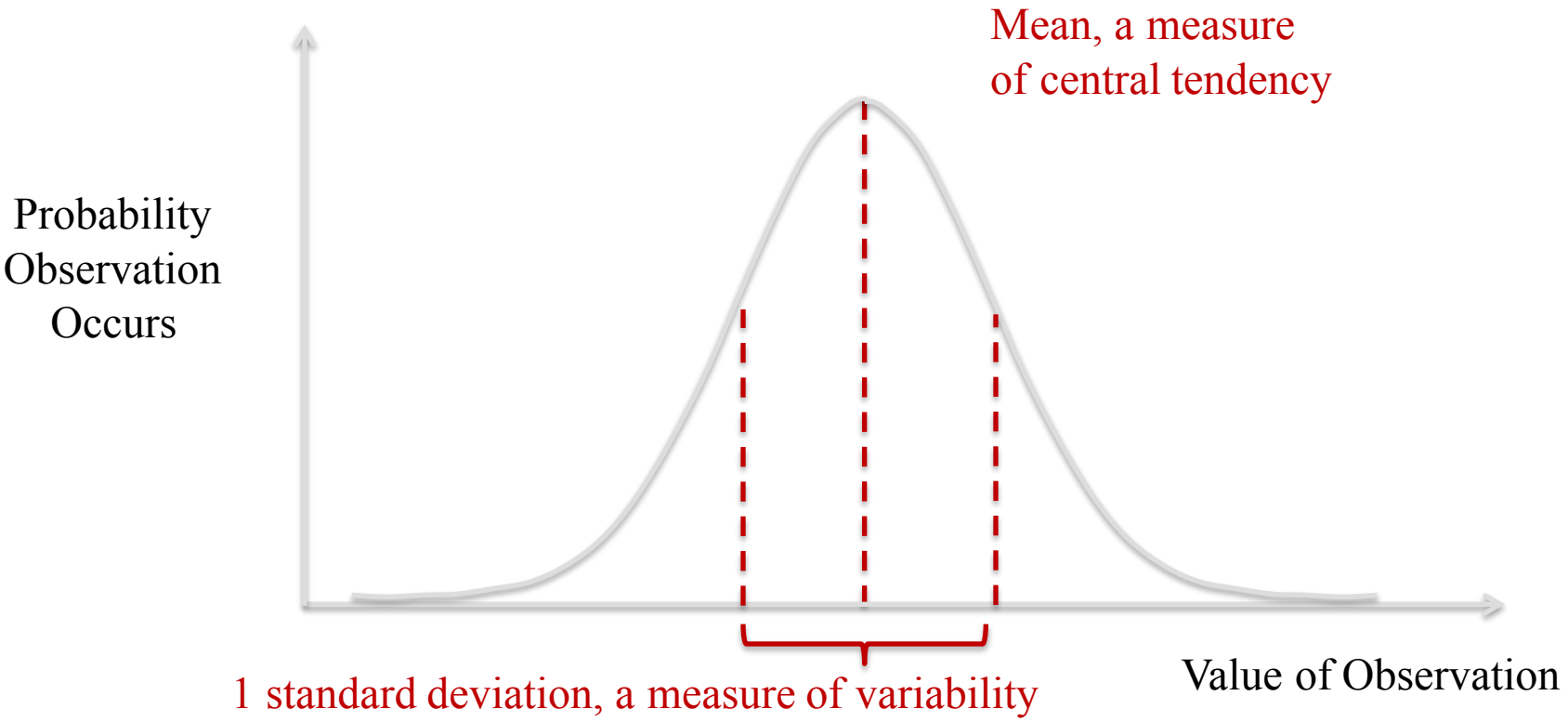
Distributions

- **Distribution: Set of all possible random variables**
- **Example:**
  - *Flipping Coins for heads and tails*
    - *a binomial distribution (two possible outcomes)*
    - *discrete (categories of heads and tails, no real numbers)*
    - *evenly weighted (heads are just as likely as tails)*
  - *Tornado events in Ann Arbor*
    - *a binomial distribution*
    - *Discrete*
    - *evenly weighted (tornadoes are rare events)*

Uniform Distribution (Continuous)

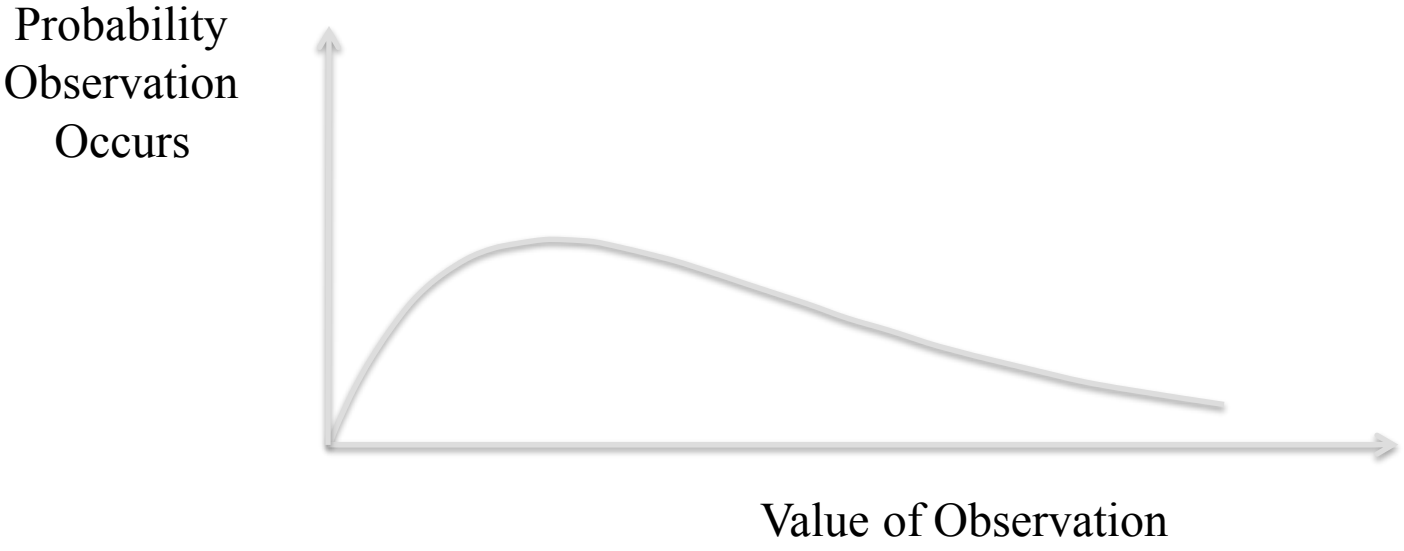


Normal (Gaussian) Distribution

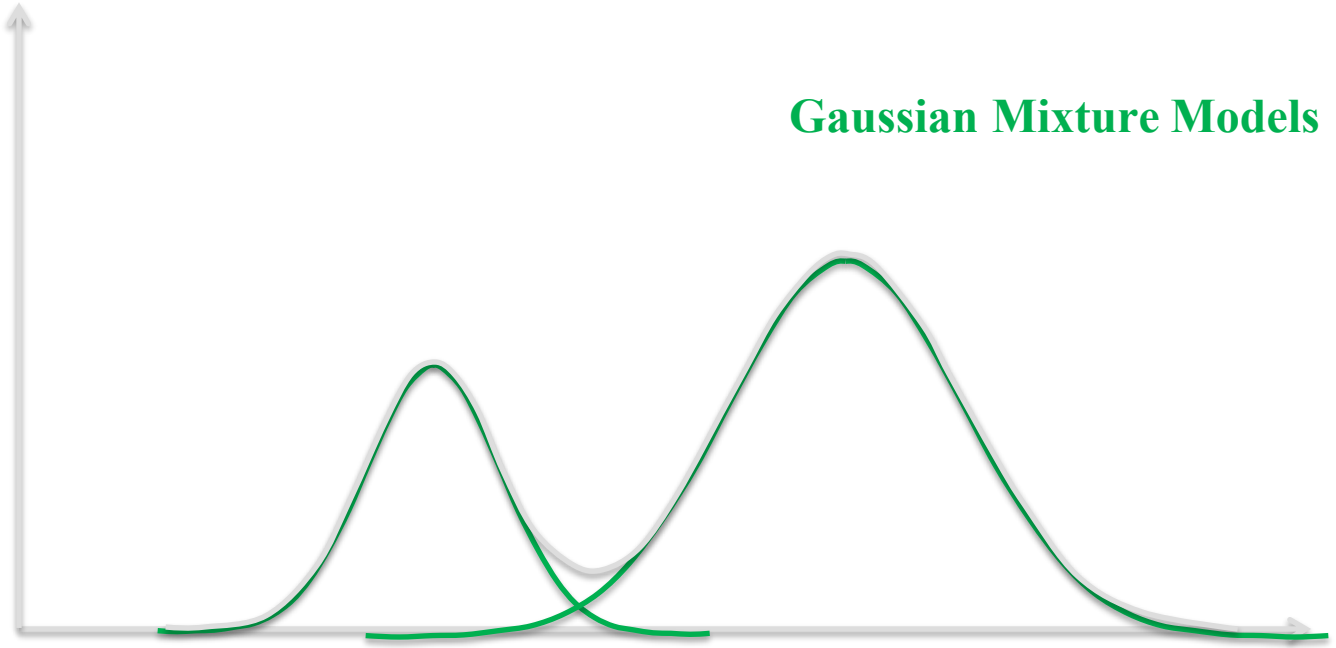


Chi Squared (χ^2) Distribution

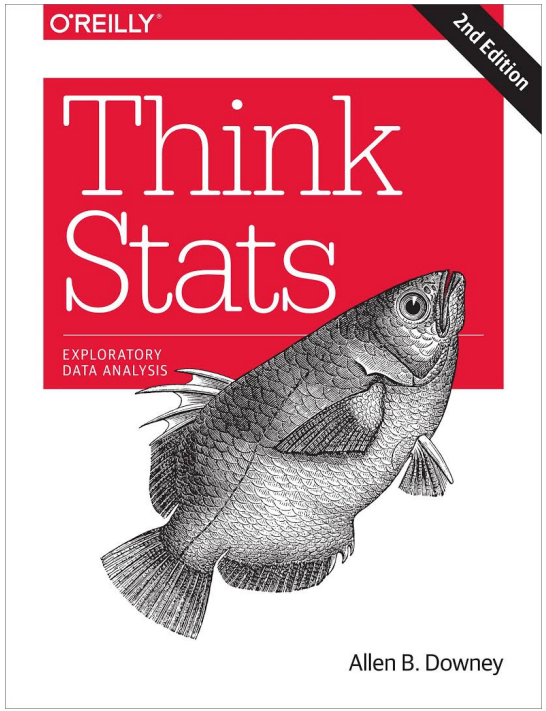
- Left-skewed
- Degrees of freedom = 4



Bimodal distributions



Think Stats



- **Probability and Statistics for Programmers**
  - *Allen B. Downey*
  - *Available for free under CC license at:*

http://greenteapress.com/thinkstats2/index.html

Hypothesis Testing

- **Hypothesis: A statement we can test**
  - *Alternative hypothesis: Our idea, e.g. there is a difference between groups*
  - *Null hypothesis: The alternative of our idea, e.g. there is no difference between groups*
- **Critical Value alpha (α)**
  - *The threshold as to how much chance you are willing to accept*
  - *Typical values in social sciences are 0.1, 0.05, or 0.01*

p-hacking

- **P-hacking, or Dredging**
  - *Doing many tests until you find one which is of statistical significance*
  - *At a confidence level of 0.05, we expect to find one positive result 1 time out of 20 tests*
  - *Remedies:*
    - *Bonferroni correction*
    - *Hold-out sets*
    - *Investigation pre-registration*