Planning developmental studies: A Bayesian Perspective

Mike Frank

How many children do I (have to) run?!?

Should this study be done?

- An alternative way to think about sample size planning: will this study be informative?
- Will your reviewers say:
  - “Likely wouldn’t have been able to reject the null no matter what”?
  - “Not precise enough to constrain future work”?
  - ”Wasted participants’ time”?

Outline

1. The (flawed) classic approach: Power analysis
2. General alternative strategies
3. How Bayesian methods can help: Sequential testing

Outline

1. **The (flawed) classic approach: Power analysis**
2. General alternative strategies
3. How Bayesian methods can help: Sequential testing

Classic NHST

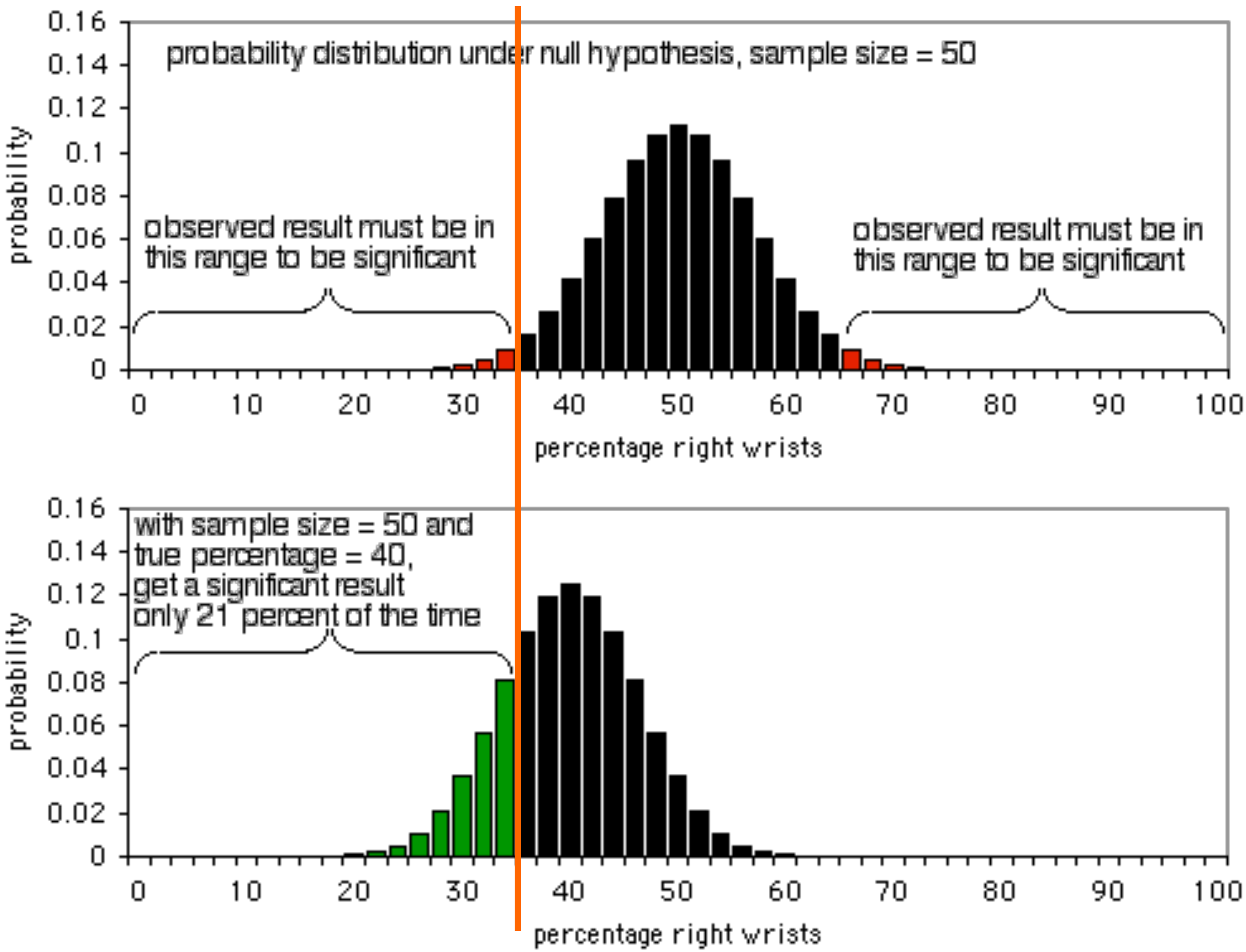
		Truth: Null is true	Truth: Null is false
Experimental result	Null is false	**Type I error** False positive	Correct
	Null is true	Correct	**Type II error** False negative

Remember: p val is the probability of the data (or any more extreme) under the null

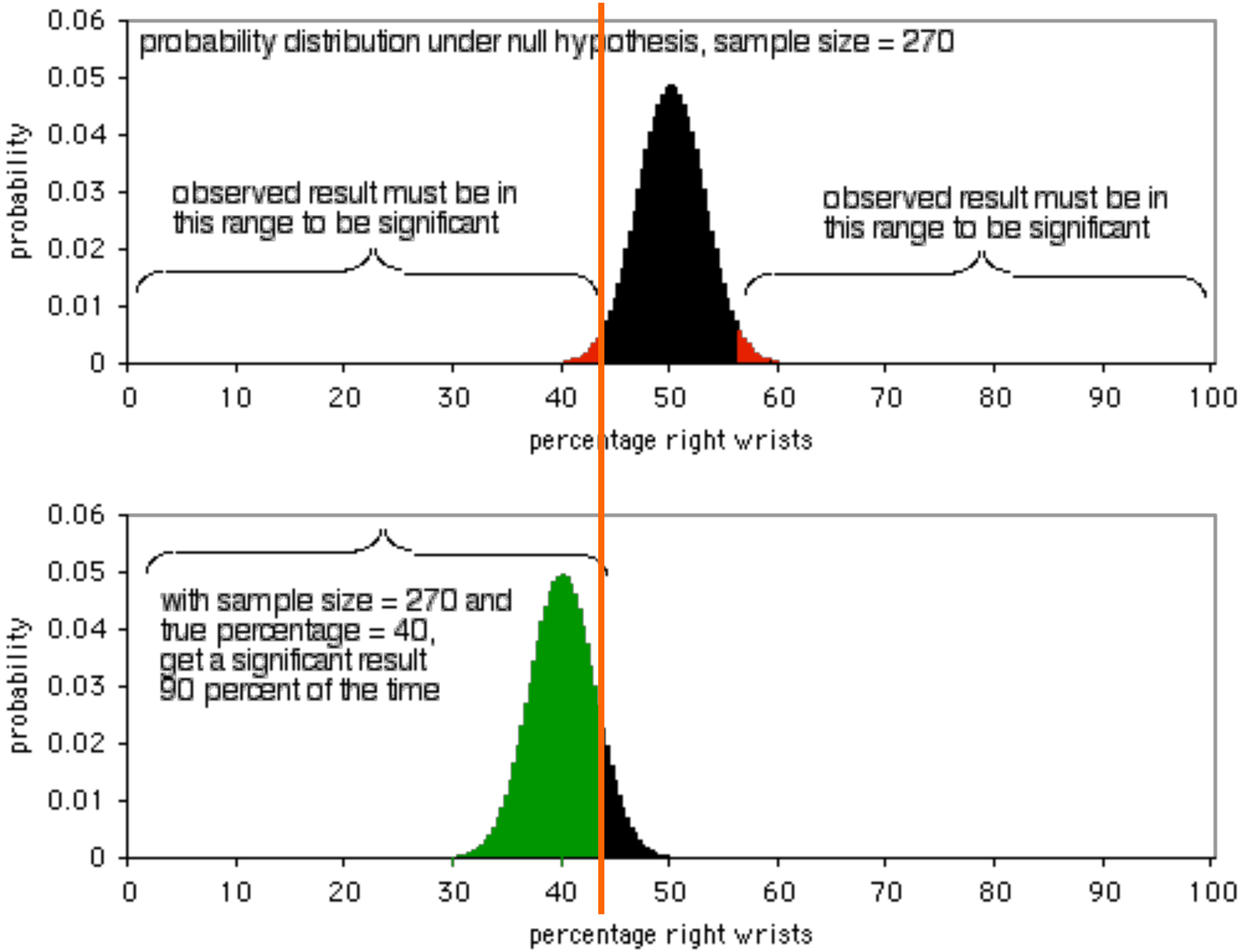
Power under classic NHST

- $\boldsymbol{\alpha}$ is the significance value
  - Also the false-positive rate!
  - Generally $\boldsymbol{\alpha} < .05$
- $\boldsymbol{\beta}$ is probability of not rejecting null
  - False negative rate
  - Power is 1- $\boldsymbol{\beta}$
- Power: The conditional probability
  - that one will reject the null hypothesis
  - given that the null is really false
  - And given
    - Effect size
    - Sample size

Quick illustration



Quick illustration



The big problem

We don’t know the real effect size!

The (other) big problem

The real effect size may be 0

(and if so, we want to accept the null)

Potential sources of effect sizes

1. Meta-analysis of previous literature
2. Previous finding you’re trying to replicate
3. General sense of the effect size you care about
4. Pilot data

Potential sources of effect sizes

1. **Meta-analysis of previous literature**
   – Great if you have it
   – But rare to have this and still be planning a study
   – Still subject to potential publication bias
2. Previous finding you’re trying to replicate
3. General sense of the effect size you care about
4. Pilot data

Potential sources of effect sizes

1. Meta-analysis of previous literature
2. **Previous finding you're trying to replicate**
   - Very likely to be an inflated effect
   - Can adjust for inflation (e.g. Biesanz & Shrager ms)
   - Still very unlikely to be a precise estimate
3. General sense of the effect size you care about
4. Pilot data

Potential sources of effect sizes

1. Meta-analysis of previous literature
2. Previous finding you're trying to replicate
3. **General sense of the effect size you care about**
   - Average effect is often small (d=.5)
   - Might want to do a "smallest effect size of interest" (SESOI) analysis
   - Planning for an average effect just ends up a recipe for relatively underpowered studies
4. Pilot data

Potential sources of effect sizes

1. Meta-analysis of previous literature
2. Previous finding you're trying to replicate
3. General sense of the effect size you care about
4. **Pilot data**
   - **DON'T DO THIS!**
   - Estimates of effect size for pilots are so noisy that they will do more harm than good!
   - Cf http://datacolada.org/20

Outline

1. The (flawed) classic approach: Power analysis
2. **General alternative strategies**
3. How Bayesian methods can help: Sequential testing

Planning a sample

- What’s the goal of the study?
  - To test the existence of an effect
  - To replicate prior work
  - To measure a particular effect for comparison to a model
- What are the resources available for completing the study?
  - How long will it take?
  - How much does each data point cost?
  - And what’s the opportunity cost?
- **Answers to these questions determine the right sample size planning method!**

Example 1

- RCT of educational intervention to raise math grades
  - High cost
  - High potential return on investment
  - Long timescale
- Want to know about efficacy of intervention
- Prior knowledge state:
  1. Lots of prior knowledge about interventions of this type
  2. Limited knowledge about effective size

1. Prior knowledge -> classic power analysis
2. Less knowledge -> consider power on range of effect sizes up to some **smallest effect size of interest (SESOI)**

Example 2

- Student project with convenience population
  - Low cost, mostly opportunity cost in terms of time
  - Limited prior knowledge
- Cost-based sample planning probably most appropriate
  - Can analyze expected power under these costs
- Consider sequential analysis to minimize costs

Example 3

- Test of judgment/decision-making model using neural data
  - High cost of data collection
  - No obvious null hypothesis to reject
- Consider precision analysis: calculate expected measurement precision as a function of spending on data collection

Outline

1. The (flawed) classic approach: Power analysis
2. **General alternative strategies**
3. How Bayesian methods can help: Sequential testing