Evidence-Informed Policy

POLSCI 4SS3

Winter 2024

Policy

- **Policy** is an umbrella term to describe government programs or operations at different levels
- **Examples:**
  - How long should form 57B be?
  - Should we get help from private clinics to clear surgey backlogs?
  - Should the education budget increase?
  - When should the next federal election be held?

Evidence-Informed

- Of course we want to base policy on evidence!
- But there is no *objective* evidence when it comes to human behavior
- We say evidence-*informed* because the best we can do is try to prove ourselves wrong, but we cannot *base* policy on evidence the same way medicine does

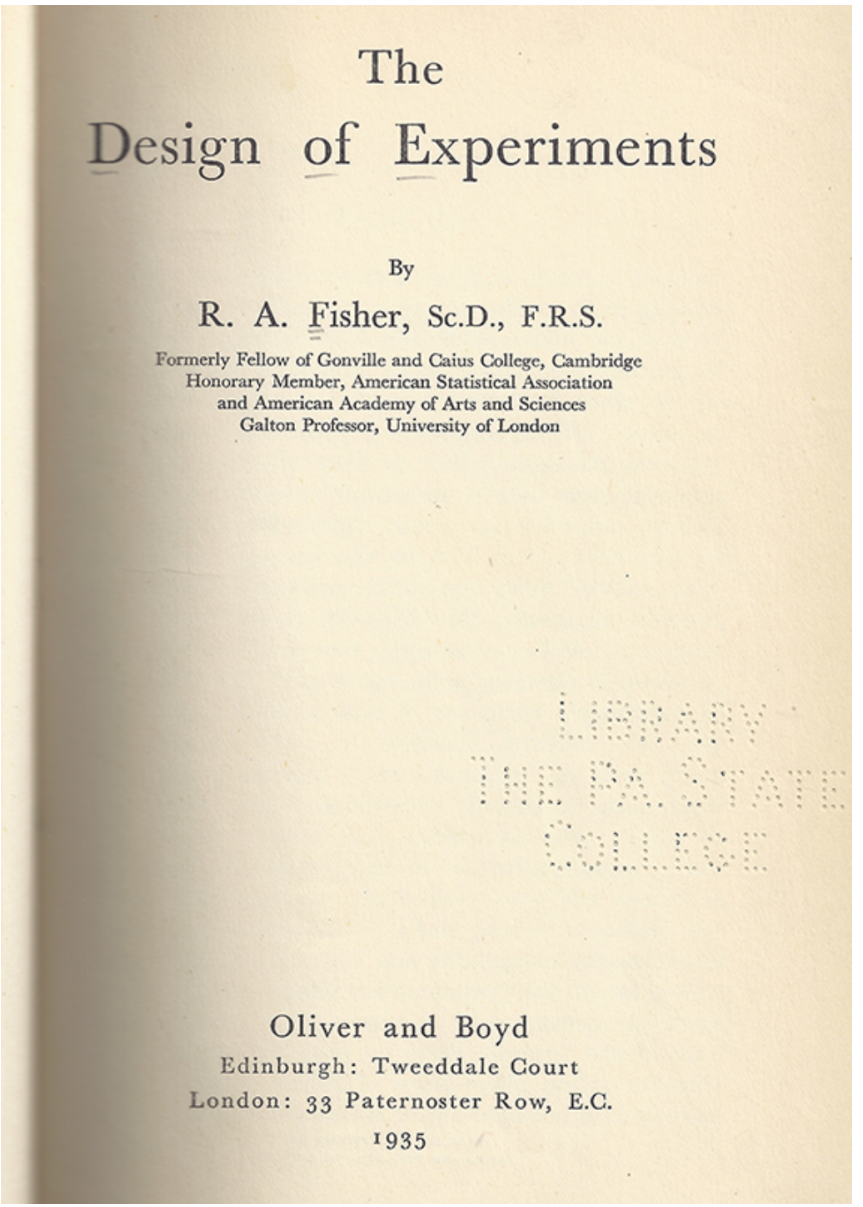
Two approaches

1. Evidence as **insight**

2. Evidence as **evaluation**

How can you determine if a policy works?

Example



The
Design of Experiments

By

R. A. Fisher, Sc.D., F.R.S.

Formerly Fellow of Gonville and Caius College, Cambridge
Honorary Member, American Statistical Association
and American Academy of Arts and Sciences
Galton Professor, University of London

Oliver and Boyd
Edinburgh: Tweeddale Court
London: 33 Paternoster Row, E.C.
1935

The lady tasting tea

A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup

How do you **evaluate** this claim?

An experiment

- Suppose we have eight milk tea cups
- 4 milk first, 4 tea first
- We arrange them in random order
- Lady knows there are 4 of each, but not which ones

Results

	True Order	
Lady's Guesses	**Tea First**	**Milk First**
Tea First	3	1
Milk First	1	3

- She gets it right $6/8$ times
- What can we conclude?

Problem

- How does “being able to discriminate” look like?
- Same for policy, we don’t know how the world where the policy works look like
- But we **do know** how a person without the ability to discriminate milk/tea order looks like
- This lets us make **probability statements** about this **hypothetical world of no effect**

A person with no ability

Count	Possible combinations	Total
0	xxxx	
1	xxxo, xxox, xoxx, oxxx	
2	xxoo, xoxo, xoox, oxox, ooxx, oxxo	
3	xooo, oxoo, ooxo, ooox	
4	oooo	

- This is symmetrical!

A person with no ability

Count	Possible combinations	Total
0	xxxx	
1	xxxo, xxox, xoxx, oxxx	
2	xxoo, xoxo, xoox, oxox, ooxx, oxxo	
3	xooo, oxoo, ooxo, ooox	
4	oooo	

A person with no ability

Count	Possible combinations	Total
0	xxxx	$1 \times 1 = 1$
1	xxxo, xxox, xoxx, oxxx	$4 \times 4 = 16$
2	xxoo, xoxo, xoox, oxox, ooxx, oxxo	$6 \times 6 = 36$
3	xooo, oxoo, ooxo, ooox	$4 \times 4 = 16$
4	oooo	$1 \times 1 = 1$

- A person just guessing gets $6/8$ cups right with probability $\frac{16}{70} \approx 0.23$

A person with no ability

Count	Possible combinations	Total
0	xxxx	$1 \times 1 = 1$
1	xxxo, xxox, xoxx, oxxx	$4 \times 4 = 16$
2	xxoo, xoxo, xoox, oxox, ooxx, oxxo	$6 \times 6 = 36$
3	xooo, oxoo, ooxo, ooox	$4 \times 4 = 16$
4	oooo	$1 \times 1 = 1$

- And **at least** $6/8$ cups with $\frac{16+1}{70} \approx 0.24$

p-values

- If the lady is **not** able to discriminate milk-tea order, the chance of observing 6/8 correct guesses or better is 24%
- We can translate this to general statements about policies or experiments
- *If* the **null hypothesis** of no effect is true…
- … the **p-value** is the probability of observing a result *equal or more extreme* than what is originally observed
- Smaller p-values give more evidence **against** the null, which helps us make a case for the policy having an effect

Diagnosing hypothesis tests

- A convention in the social sciences is to claim that something with $p < 0.05$ is *statistically significant*[1]
- Committing to a **significance level** implies accepting that sometimes we will get $p < 0.05$ by chance
- This is a **false positive** result
- A good answer strategy as a **controlled** false positive rate `(more in the lab!)`

Next Two Weeks

Field Experiments

Focus on: Research design alternatives

Break time!


Lab