Modern Sampling Methods

Class 3: Publication Bias and Pre-Analysis Plans

January 10, 2022

Outline

- Publication Bias
  - Background
  - Framework
  - Detection and Correction
  - Pre-Analysis Plans

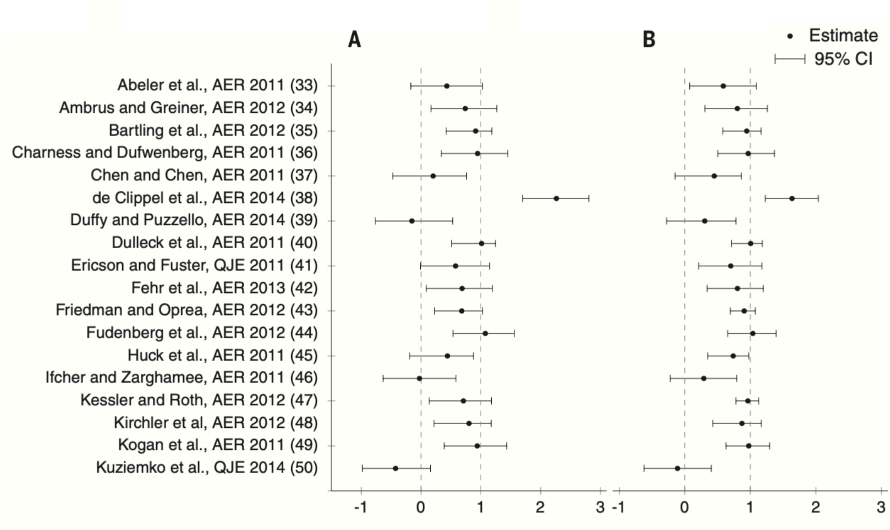
Publication Bias

There is some evidence that in fields where statistical tests of significance are commonly used, research which yields nonsignificant results is not published. Such research being unknown to other investigators may be repeated independently until eventually by chance a significant result occurs—an "error of the first kind"—and is published. Significant results published in these fields are seldom verified by independent replication. The possibility thus arises that the literature of such a field consists in substantial part of false conclusions resulting from errors of the first kind in statistical tests of significance.

From: Sterling (1959)

- Publication Gatekeepers
- Researcher Behavior - "File Drawer Problem"

Replication Results



From: Camerer et al (2016)

Framework (Andrews-Kasy 2019)

- Stylized model of publication bias
- Each study focuses on learning about a true effect, Θ^*
- Researcher has access to data which is used to produce a "noisy" estimate, $\widehat{\Theta}^*$
- Noise is summarized by a standard error, Σ^*
- Test statistic for statistical significance is $Z^* = \frac{\widehat{\Theta}^*}{\Sigma^*}$
- Publication is a determined by a draw from a Bernoulli distribution with probabilty $p(Z^*)$

$(\Theta^*, \Sigma^*, \widehat{\Theta}^*, D)$

$$\begin{array}{lrcl}
\text{True Effect , Std Err:} & (\Theta^*, \Sigma^*) & \sim & \mu_{\Theta,\Sigma} \\
\text{Result / Estimate:} & \widehat{\Theta}^* \mid \Theta^*, \Sigma^* & \sim & N(\Theta^*, \Sigma^{*2}) \\
\text{Publication Indicator:} & D \mid \widehat{\Theta}^*, \Theta^*, \Sigma^* & \sim & \text{Bernoulli}(p(Z^*))
\end{array}$$

where

- $Z^* = \frac{\widehat{\Theta}^*}{\Sigma^*}$ and $\Omega^* = \frac{\Theta^*}{\Sigma^*}$
- Superscript * denotes latent process
- $D = \begin{cases} 1 & \text{Observed / Published} \\ 0 & \text{Unobserved / Unpublished} \end{cases}$
- e.g. $p(Z^*) = \begin{cases} .75 & \text{if } |Z^*| \geq 1.96 \\ 0.1 & \text{if } |Z^*| < 1.96 \end{cases}$

Selection / Bias

- The publication probability $p(\cdot)$ creates a wedge between the latent and observed distributions
- Conditional distribution of latent (normalized) estimator

$$f_{Z^* \,|\, \Omega^*, \Sigma^*}(z \,|\, \omega, \sigma) \;=\; \varphi(z - \omega)$$

  - $\widehat{\Theta}^* \,\big|\, \Theta^* = \theta, \Sigma^* = \sigma \;\sim\; N(\theta, \sigma^2)$
  - Let $\omega = \frac{\theta}{\sigma}$. Then, $Z^* = \frac{\widehat{\Theta}^*}{\Sigma^*} \,\big|\, \Omega^* = \omega, \Sigma^* = \sigma \;\sim\; N(\omega, 1)$

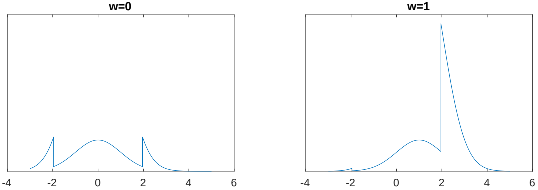
- Conditional distribution of observed (normalized) estimator

$$f_{Z \,|\, \Omega, \Sigma}(z \,|\, \omega, \sigma) \;\propto\; p(z)\varphi(z - \omega)$$

  - Let Z denote observed Z^* (when $D = 1$)
  - $Z \,\big|\, \Omega = \omega, \Sigma = \sigma \;\sim\; Z^* \,\big|\, \Omega^* = \omega, \Sigma^* = \sigma, D = 1$
  - Apply Bayes Rule

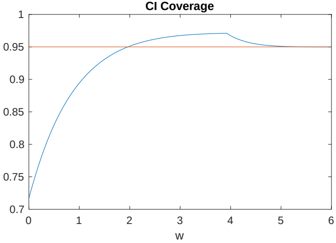
Distribution of observed (normalized) estimator

$Z \mid \Omega = \omega$



Confidence Interval Coverage

- $\text{CI} = [Z - 1.96, Z + 1.96]$
- $\text{Coverage} = \Pr(Z - 1.96 \leq \omega \leq Z + 1.96 \mid \Omega = \omega)$



Tests and Corrections

In this stylized framework, $p(\cdot)$ determines the degree of publication bias.

So, given $p(\cdot)$ or $\hat{p}(\cdot)$, one could

- Correct Confidence Intervals
  Use median unbiased adjustment of Z to center CI
  Adjust critical values to achieve correct coverage
- Check / Test for publication bias
  $H_0 : p(\cdot) = \text{constant}$

Identification / Estimation of $p(\cdot)$,

- Replication Studies
- Meta Studies

Replication Studies

$(\Theta^{R*}, \Sigma^{R*}, \widehat{\Theta}^{R*}, D^R)$

Assume:

- Same effect: $\Theta^{R*} = \Theta^*$
- Pub bias through initial study:
  $D^R = D$ which depends on Z^* but not Z^{R*}
- $$\widehat{\Theta}^{R*} \,\big|\, \Theta^*, \Sigma^{R*}, \Sigma^*, D, \widehat{\Theta}^* \;\sim\; N(\Theta^*, (\Sigma^{R*})^2)$$

$\Rightarrow$ Identification of $p(\cdot)$ from asymmetries in (Z^R, Z) distribution

Andrews-Kasy (2019) estimate based on Camerer (2016) replication data:

- $\hat{p}(z) = \begin{cases} 1 & \text{if } |z| \geq 1.96 \\ 0.029 & \text{if } |z| < 1.96 \end{cases}$
- Reject null of no publication bias

Meta Regressions

Assume:

- Independence: $\Theta^* \perp \Sigma^*$

$\Rightarrow$ Identification of $p(\cdot)$ follows from identification of μ_Θ using variation in $\frac{f_{Z|\Sigma}(z|\sigma_2)}{f_{Z|\Sigma}(z|\sigma_1)}$

Common Regressions:

$$\widehat{\Theta} = \gamma_0 + \gamma_1 \Sigma + \varepsilon \quad \text{and} \quad Z = \lambda_0 + \lambda_1 \frac{1}{\Sigma} + \eta$$

Under $H_0 : p(\cdot) = \text{constant}$ (no pub bias),

$$\gamma_0 = E[\Theta^*],\ \gamma_1 = 0 \quad \text{and} \quad \lambda_0 = 0,\ \lambda_1 = E[\Theta^*]$$

So, $\gamma_1 = 0$ or $\lambda_0 = 0$ used to test for publication bias

Researcher Process

- Small extension of the framework to consider the role of pre-analysis plans
  - Allow for multiple (possibly dependent) effects in a given study
  - Include a new step connecting data to empirical results
- Let W^* denote data, and

$$d = \begin{cases} 1 & \text{if } Z^* \text{ is reported by researcher} \\ 0 & \text{if } Z^* \text{ is unreported by researcher} \end{cases}$$

- Extended framework

$$\begin{aligned}
(i) \quad & (\Theta_1^*, \Sigma_1^*), \ldots, (\Theta_N^*, \Sigma_N^*) \sim \mu_{\Theta,\Sigma} \\
(ii) \quad & (W_1^*, \ldots, W_N^*) \,|\, (\Theta_1^*, \Sigma_1^*), \ldots, (\Theta_N^*, \Sigma_N^*) \sim \mu_{W|\Theta,\Sigma} \\
(iii) \quad & \widehat{\Theta}_j^* = g_j(W_j^*) \\
(iv) \quad & (d_1, \ldots, d_N) = h(Z_1^*, \ldots, Z_N^*)
\end{aligned}$$

TABLE VI

ERRONEOUS INTERPRETATIONS UNDER "CHERRY PICKING"

Outcome variable	(1) Mean for controls	(2) Treatment effect
Panel A: GoBifo "weakened" institutions		
Attended meeting to decide what to do with the tarp	0.81	-0.04^{+}
Everybody had equal say in deciding how to use the tarp	0.51	-0.11^{+}
Community used the tarp (verified by physical assessment)	0.90	-0.08^{+}
Community can show research team the tarp	0.84	–0.12*
Respondent would like to be a member of the VDC	0.36	–0.04*
Respondent voted in the local government election (2008)	0.85	–0.04*
Panel B: GoBifo "strengthened" institutions		
Community teachers have been trained	0.47	0.12^{+}
Respondent is a member of a women's group	0.24	0.06**
Someone took minutes at the most recent community meeting	0.30	0.14*
Building materials stored in a public place when not in use	0.13	0.25*
Chiefdom official did not have the most influence over tarp use	0.54	0.06*
Respondent agrees with "Responsible young people can be good leaders" and not "Only older people are mature enough to be leaders"	0.76	0.04*
Correctly able to name the year of the next general elections	0.19	0.04*

From: Casey, Glennerster, and Miguel (2012)

Pre-registration / Pre-Analysis Plans

- Describes (i) and (ii) and hypotheses to be considered (Registration)
- Specifies g_j for $j = 1, \ldots, N$ and any tests, ... (Analysis)
- Commits to $d_j = 1 \; \forall j$ (Registration/Analysis)

$$
\begin{array}{ll}
(i) & (\Theta_1^*, \Sigma_1^*), \ldots, (\Theta_N^*, \Sigma_N^*) \sim \mu_{\Theta,\Sigma} \\
(ii) & (W_1^*, \ldots, W_N^*) \,|\, (\Theta_1^*, \Sigma_1^*), \ldots, (\Theta_N^*, \Sigma_N^*) \sim \mu_{W|\Theta,\Sigma} \\
(iii) & \widehat{\Theta}_j^* = g_j(W_j^*) \\
(iv) & (d_1, \ldots, d_N) = h(Z_1^*, \ldots, Z_N^*)
\end{array}
$$

Pre-Analysis Plans

Table 1
Pre-Analysis Plan Checklist

Item	*Brief description*
Primary outcome variable	The key variable of interest for the study. If multiple variables are to be examined, one should know how the multiple hypothesis testing will be done.
Secondary outcome variable(s)	Additional variables of interest to be examined.
Variable definitions	Precise variable definitions that specify how the raw data will be transformed into the actual variables to be used for analysis.
Inclusion/Exclusion rules	Rules for including or excluding observations, and procedures for dealing with missing data.
Statistical model specification	Specification of the precise statistical model to be used, hypothesis tests to be run.
Covariates	List of any covariates to be included in analysis.
Subgroup analysis	Description of any heterogeneity analysis to be performed on the data.
Other issues	Other issues include data monitoring plans, stopping rules, and interim looks at the data.

From: Olken (2015)

Overview

- Clinical Trials / FDA
- AEA RCT Registry
- Recommendations / Trade-offs
  - Banerjee et al (2020)
  - Coffman and Niederle (2015)
  - FDA (1998)
  - Finkelstein et al (2020)
  - Ludwig, Mullainathan, Speiss (2019)
  - Miguel (2021)
  - Olken (2015)