Which findings get published?
Which findings should be published?

Maximilian Kasy

December 10, 2018

Introduction

- Replicability is a fundamental requirement of science. Different researchers should reach the same conclusions. Methodological conventions should ensure this.
- Replications of published experiments frequently find effects which are of smaller magnitude or opposite sign.
- One explanation: Selective publication based on findings.
  1. Publication bias
     - Journal editor and referee decisions.
     - Statistical significance, surprisingness, or confirmation of prior beliefs.
  2. P-hacking and specification searching
     - Researcher decisions.
     - Incentives to select which findings to submit based on the likelihood of publication.

Two questions

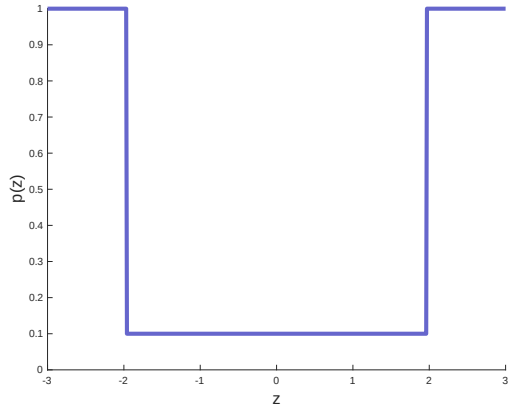
1. Which findings get published?
   - How much and based on what criteria are findings selected?
   - How can we correct for such selection?
   - Existing approaches test whether publication is selective, but do not estimate the amount and form of selection.
2. Which findings should be published?
   - Replicability is not the only goal of research.
   - Relevance for policy (and other) decisions is important, as well.
   - These two goals might potentially stand in conflict.
   - Existing reform proposals focus on replicability and aim to eliminate selection, ignoring the role of relevance.

Andrews, I. and Kasy, M. (2018). Identification of and correction for publication bias

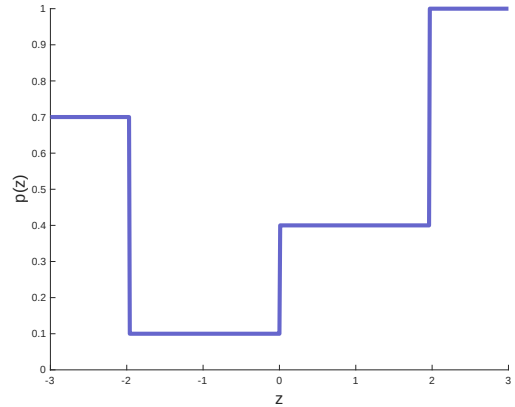
Frankel, A. and Kasy, M. (2018). Which findings should be published?

Examples: Possible forms of selection $p(Z)$

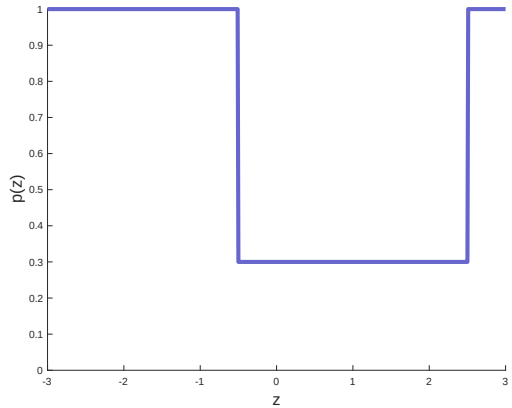
Significance



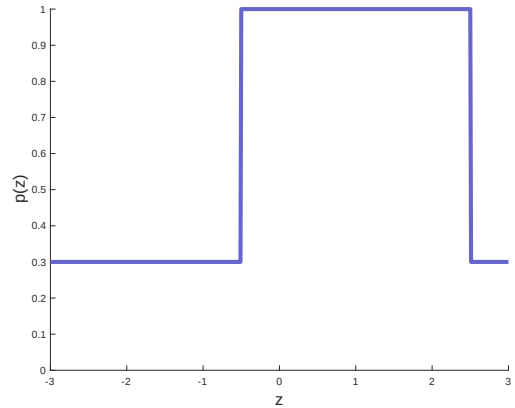
Significance and sign



Surprisingness



Plausibility



- $p(Z)$: Probability that an estimate Z is published.

Question 1: Which findings get published?

Key results

1. If form and magnitude of selection are known, we **can correct published findings**.
   - Unbiased estimates, confidence sets that control size.
   - Using "quantile inversion."
2. Form and magnitude of **selection** are **nonparametrically identified**.
   - Using systematic replication studies. Absent selectivity, original and replication estimates should be symmetrically distributed.
   - Using meta-studies. Absent selectivity, distribution of estimates for small sample sizes should be noised-up version of distribution for larger sample sizes.
3. **Published research is selected**:
   - Lab experiments in economics and psychology:
     Statistical significance
   - Effect of minimum wages on employment:
     Statistical significance, sign.
   - Deworming:
     Inconclusive.

Question 2: Which findings should be published?

Reforming scientific publishing

- Publication bias motivates calls for reform:
  Publication should not select on findings.
  - De-emphasize statistical significance, ban "stars."
  - Pre-analysis plans to avoid selective reporting of findings.
  - Registered reports reviewed and accepted prior to data collection.
- But: Is eliminating bias the right objective?
  How does it relate to informing decision makers?
- We characterize **optimal publication rules from an instrumental perspective**:
  - Study might inform the public about some state of the world.
  - Then the public chooses a policy action.
  - Take as given that not all findings get published (prominently).

Which findings should be published?

Key results

1. **Optimal** rules selectively **publish surprising findings**.
   In leading examples: Similar to two-sided or one sided tests.
2. But: Selective publication **always distorts inference**.
   There is a trade-off policy relevance vs. statistical credibility.
3. With **dynamics**: Additionally publish **precise null** results.
4. With **incentives**: Modify publication rule to **encourage more precise** studies.

Setup

Timeline and notation

State of the world (parameter)	θ
Common prior	$\theta \sim \pi_0$
Study might be submitted	
Exogenous submission probability	q
Design (e.g., standard error)	$S \perp \theta$
Findings (estimate)	$X \vert \theta, S^2 \sim f_{X \vert \theta, S}$
Journal decides whether to publish	$D \in \{0, 1\}$
Publication probability	$p(X, S)$
Publication cost	c
Public updates beliefs	$\pi_1 = \pi_1^{(X,S)}$ if $D = 1$
	$\pi_1 = \pi_1^0$ if $D = 0$
Public chooses policy action	$a = a^*(\pi_1) \in \mathbb{R}$
Utility	$U(a, \theta)$
Social welfare	$U(a, \theta) - Dc$

Optimal publication rules

- We show that

ex-ante optimal rules, maximizing expected welfare, are those which
ex-post publish findings that have a big impact on policy.

- **Interim gross benefit** $\Delta(\pi, a^0)$ of publishing equals
  - Expected welfare given publication, $\mathbb{E}_{\theta\sim\pi}[U(a^*(\pi), \theta)]$,
  - minus expected welfare of default action, $\mathbb{E}_{\theta\sim\pi}[U(a^0, \theta)]$.
- **Interim optimal publication rule**:
  Publish if interim benefit exceeds cost c.
- Want to maximize **ex-ante expected welfare**:

$$\begin{aligned} EW(p, a^0) =& \mathbb{E}[U(a^0, \theta)] \\ &+ q \cdot \mathbb{E}\Big[p(X, S) \cdot (\Delta(\pi_1^{(X,S)}, a^0) - c)\Big]. \end{aligned}$$

- Immediate consequence:
  Optimal policy is interim optimal given a^0.

Two key results

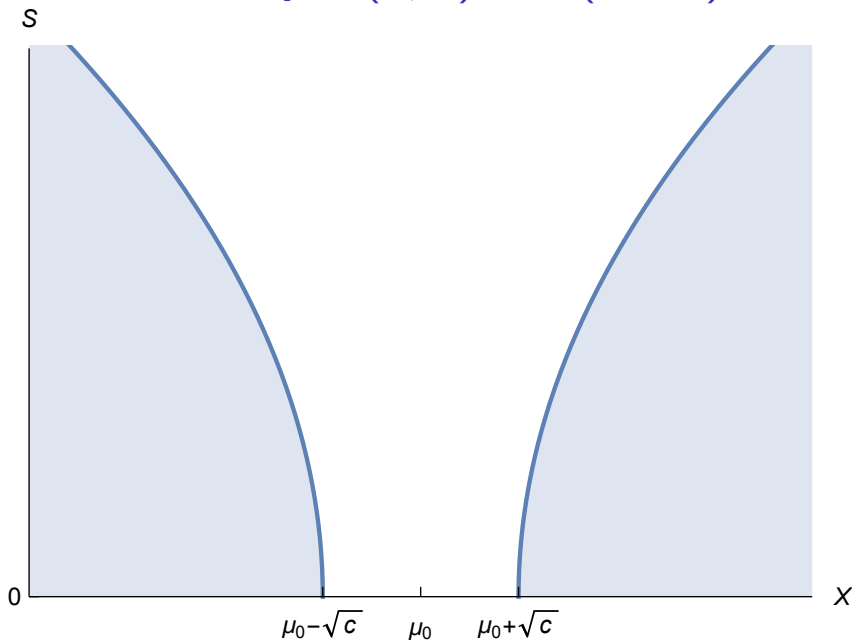
- **Don't publish null results:**
  A finding that induces $a^*(\pi^I) = a^0 = a^*(\pi_1^0)$ always has 0 interim benefit and should never get published.

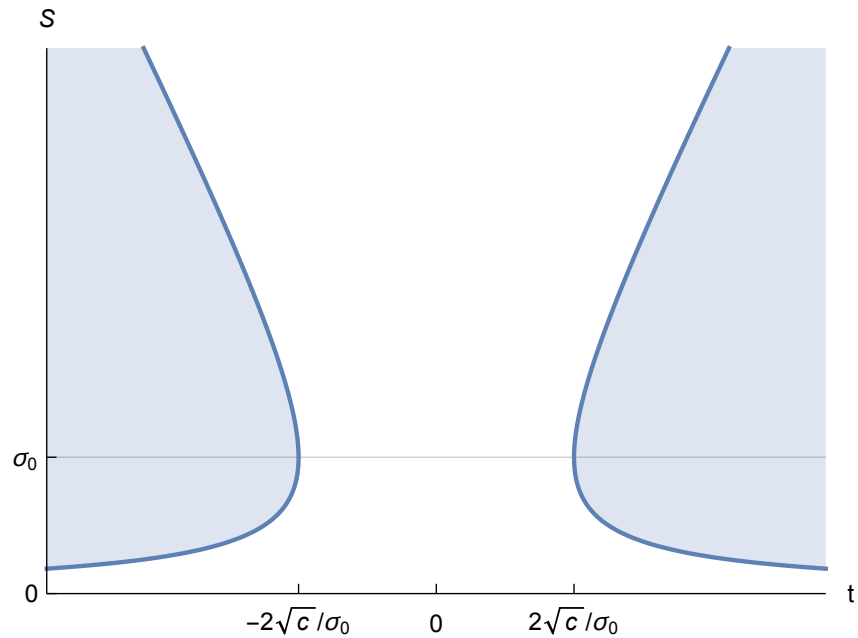
- **Publish findings outside interval:**
  Suppose
  - U is supermodular.
  - $f_{X|\theta,S}$ satisfies monotone likelihood ratio property given $S = s$.
  - Updating is either naive or Bayes.

  Then there exists an interval $I^s \subseteq \mathbb{R}$ such that (X, s) is published under the optimal rule if and only if $X \notin I^s$.

Leading examples

Quadratic loss utility $U(a,\theta) = -(a-\theta)^2$, normal prior, normal signals

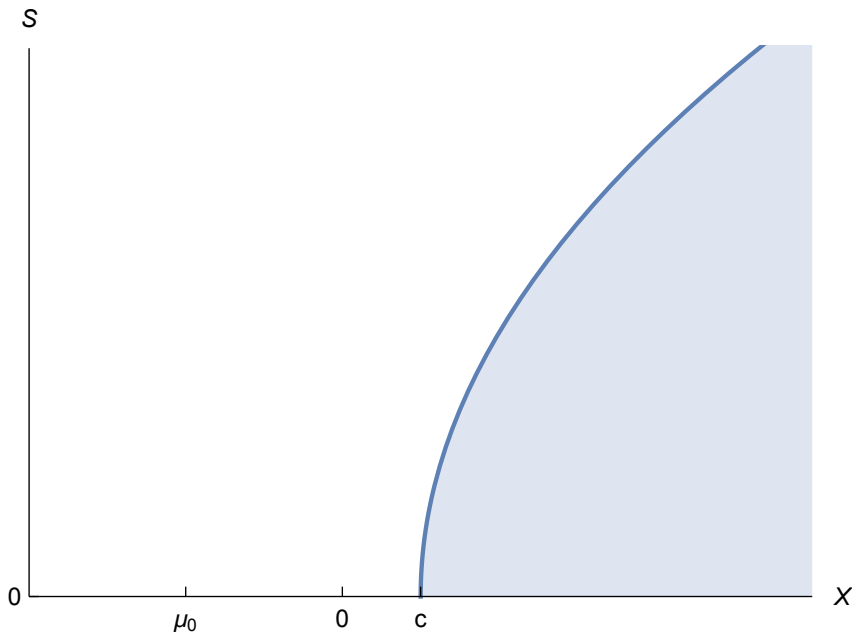


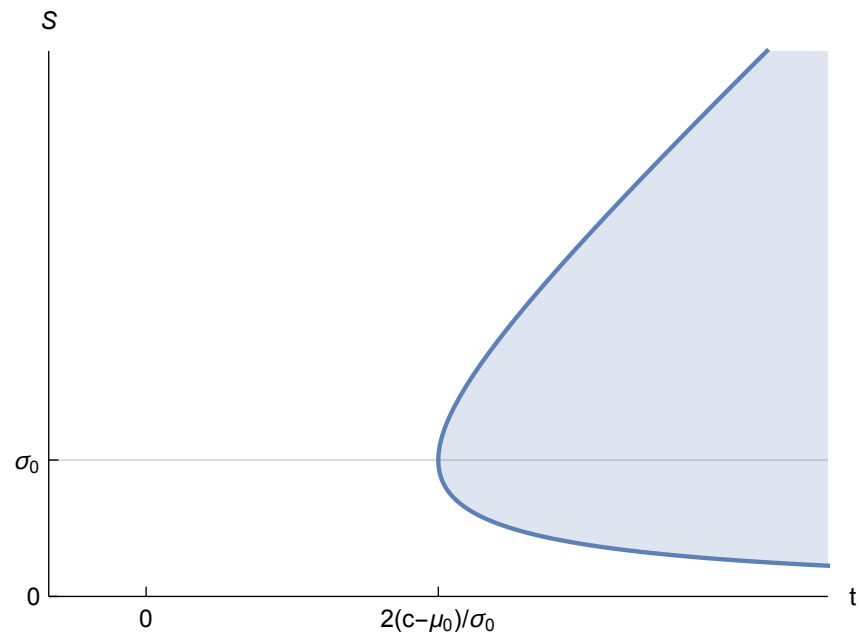


- Optimal publication region (shaded). Axes:
  - left Estimate X, standard error S.
    (As in meta-studies plots!)
  - right "t-statistic" $t = (X-\mu_0)/S$, standard error S.
- Note:
  - Given S, publish outside symmetric interval around μ_0.
  - Critical value for t-statistic is non-monotonic in S.

Leading examples

Binary action utility $U(a, \theta) = a \cdot \theta$, normal prior, normal signals





- Optimal publication region (shaded). Axes:
  - left Estimate X, standard error S.
  - right "t-statistic" $t = (X - \mu_0)/S$, standard error S.
- Note:
  - When prior mean is negative, optimal rule publishes for large enough positive X.

Outlook

Different ways of thinking about statistics / econometrics:

1. Making decisions based on data.
   - Objective function?
   - Set of feasible actions?
   - Prior information?
2. Statistics as (optimal) communication.
   - Not just "you and the data."
   - What do we communicate to whom?
   - Subject to what costs and benefits?
     Why not publish everything? Attention?
3. Statistics / research as a social process.
   - Researchers, editors and referees, policymakers.
   - Incentives, information, strategic behavior.
   - Social learning, paradigm changes.

Much to be done!

A web-app for estimating publication bias in meta-studies is available at

`https://maxkasy.github.io/home/metastudy/`

Thank you!