Pre-analysis plans and mechanism design

Maximilian Kasy Jann Spiess

January 2020

Introduction

- Trial registration and pre-analysis plans (PAPs) have become a standard requirement for experimental research.
  - For clinical studies in medicine starting in the 1990s.
  - For experimental research in economics more recently.
- Standard justification: Guarantee validity of inference.
  - P-hacking, specification searching, and selective publication distort inference.
  - Tying researchers' hands prevents selective reporting.
  - "PAPs are to frequentist inference what RCTs are to causality."
- Counter-arguments:
  - Pre-specification is costly.
  - Interesting findings are unexpected and flexibility is necessary.

Open questions

1. Why do we need a commitment device?
   Standard decision theory has no time inconsistency!

2. Under what conditions are PAPs more or less useful?
   How do we trade off the benefits and costs of PAPs?

Our approach

- Import insights from contract theory / mechanism design to statistics.
  - PAPs can be rationalized with multiple parties, conflicts of interest, and costly communication / asymmetric information.
  - We consider (optimal) statistical decision rules subject to the constraint of implementability.

- Our model:
  1. A journal commits to a publication / testing rule,
  2. then a researcher commits to a PAP,
  3. then observes the data, reports selected statistics to the journal,
  4. which then applies the publication / testing rule.

- PAPs are optimal when
  - there are many researcher degrees of freedom,
  - and/or communication costs are high.

Literature

- **P-hacking and publication bias**
  Ioannidis (2005), Gelman and Loken (2013), Andrews and Kasy (2019)
- **Contract theory and mechanism design**
  Hurwicz (1972), Mas-Colell et al. (1995) chapter 23.
- **Discussions of PAPs by empirical practitioners**
  Food and Drug Administration (1998), Coffman and Niederle (2015), Olken (2015), Christensen and Miguel (2016), Duflo et al. (2020)
- **Applied theory of the publication process**
  Ottaviani and Squintani (2006), Frankel and Kasy (2018), Spiess (2018)

Introduction

Baseline model

- Assumptions
- Implementability and optimality

Analysis

- A minimal example: $\bar{n} = 3$
- Symmetric publication rules
- General solution

Model variations

- Frequentist testing
- Multiple parameters / hypotheses
- Cost of observation

Conclusion

Setup

- Two agents: Researcher and journal.
- The researcher observes a vector

$$X = (X_1, \ldots, X_{\bar{n}}),$$

  where

$$X_i \overset{\text{iid}}{\sim} \text{Ber}(\theta).$$

- Researcher: Reports a subvector X_I to the journal, where

$$I \subset \{1, \ldots, \bar{n}\}.$$

- Journal: Makes a decision

$$a \in \{0, 1\},$$

  based on this report.

Prior and objectives

- Common prior:
$$\theta \sim \text{Beta}(\alpha, \beta).$$

- Researcher's objective:
$$u^{res} = a - c \cdot |I|.$$

  $|I|$ is the size of the reported set,
  c is the cost of communicating an additional component.

- Journal's objective:
$$u^{jour} = a \cdot (\theta - \underline{\theta}).$$

  $\underline{\theta}$ is a commonly known parameter.
  Minimum value of θ beyond which the journal would like to choose $a = 1$.

Timeline

1. The journal commits to a publication rule

$$a = a(J, I, X_I).$$

2. The researcher reports a PAP

$$J \subseteq \{1, \ldots, \bar{n}\}.$$

3. The researcher next observes X, chooses $I \subseteq \{1, \ldots, \bar{n}\}$, and reports

$$(I, X_I).$$

4. The publication rule is applied and utilities are realized.

Alternative interpretations of this stylized model

1. Publication decision:
   - A researcher wants to get published.
   - A journal wants to publish only studies for large enough true effects.
2. Drug approval:
   - A pharma company wants drug approval.
   - The public authority (FDA) wants to approve only effective drugs.
3. Hypothesis testing:
   - A researcher wants to reject the null (always).
   - A reader wants to only reject when $\theta > \underline{\theta}$.

Implementability

- Let x denote values that the random vector X may take.
- Reduced form mapping (statistical decision rule)

$$x \to \bar{a}(x).$$

- $\bar{a}(x)$ is implementable
if there exist mappings $I(x)$ and $a(I, x_I)$
such that for all x

$$\bar{a}(x) = a(I(x), X_{I(x)}),$$

and

$$I(x) \in \underset{I}{\operatorname{argmax}} \; a(I, x_I) - c \cdot |I|.$$

Optimal implementable publication rules

- The latter is the incentive compatibility constraint, which implies
  1. $$I(x) \in \underset{I}{\operatorname{argmin}} \{|I| : \ a(I, x_I) = 1\}$$
     whenever $\bar{a}(x) = 1$, and $I(x) = \emptyset$ else.
  2. $$|I(x)| \leq 1/c$$
     for all x.
- Our agenda:
  - Find implementable mappings (decision rules) $\bar{a}(x)$
  - that maximize the expected journal utility $E[u^{jour}]$.

Notation

- Successes among all components: $S = \sum_{i=1}^{\bar{n}} X_i$.
  Successes among the subset I: $S_I = \sum_{i \in I} X_i$.

- Maximal number of components the researcher is willing to submit:

$$\bar{n}^{PC} = \max\left\{n: \ 1 - cn \geq 0\right\} = \lfloor 1/c \rfloor .$$

- First-best publication cutoff for the journal:

$$\underline{s}^*(n) = \min\left\{\underline{s}: \ E[\theta \middle| S_{1,\ldots,n} = \underline{s}] \geq \underline{\theta}\right\} .$$

Introduction

Baseline model
- Assumptions
- Implementability and optimality

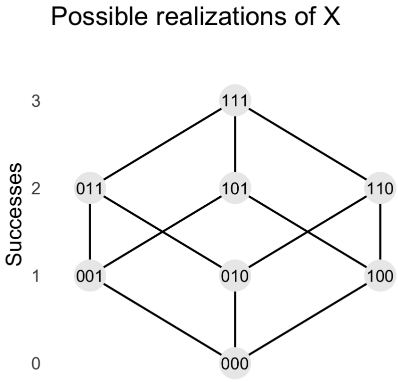
Analysis
- A minimal example: $\bar{n} = 3$
- Symmetric publication rules
- General solution

Model variations
- Frequentist testing
- Multiple parameters / hypotheses
- Cost of observation

Conclusion

A minimal example: $\bar{n} = 3$

- Suppose $\bar{n} = 3$.
  Possible realizations of X form a cube.
- Suppose $\bar{n}^{PC} = 2$.
  Possible reports $(I, X_I) \approx$
  edges of the cube.
- Reduced form mappings $\bar{a}(x) \approx$
  set of nodes for which $a = 1$.
- Vertical axis $=$
  number of successes S.



A minimal example: $\bar{n} = 3$
Case I: Symmetric cutoff rule is optimal

- Suppose $\bar{n} = 3$, $\bar{n}^{PC} = 2$, and $\underline{s}^*(3) = 2$.
- The **unconstrained efficient** solution is given by

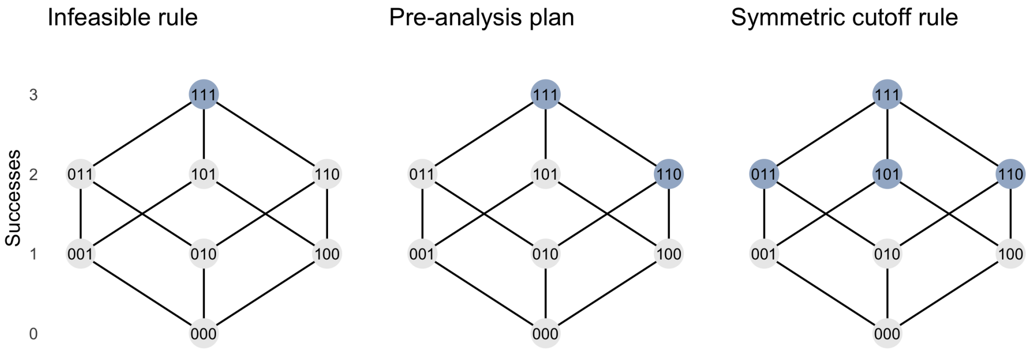
$$\bar{a}(X) = \mathbf{1}(S \geq 2).$$

- This solution **can be implemented** by

$$a(I, X_I) = \mathbf{1}(S_I \geq 2).$$

- **No PAP** is needed to implement this solution.

A minimal example: $\bar{n} = 3$



A minimal example: $\bar{n} = 3$
Case II: PAP is optimal

- Suppose again that $\bar{n} = 3$, and $\bar{n}^{PC} = 2$. Suppose now

$$\underline{s}^*(3) = 3, \qquad \underline{s}^*(2) = 2$$

- The **unconstrained efficient** solution is given by

$$\bar{a}(X) = \mathbf{1}(S = 3).$$

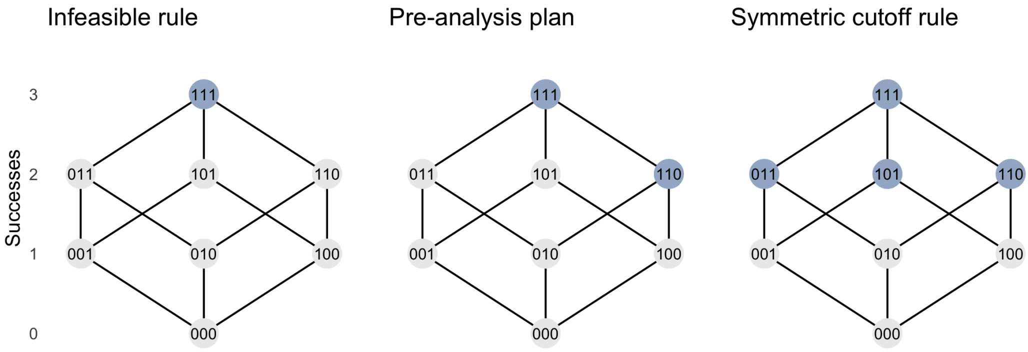
  There is **no** incentive compatible **implementation** of this solution.

- The **PAP** solution for $J = \{1, 2\}$,

$$a(J, I, X_I) = \mathbf{1}(I = \{1, 2\}, S_I = 2),$$

  yields $E[u^{jour}] > 0$, and is **constrained optimal**.

A minimal example: $\bar{n} = 3$



Symmetric publication rules

- Denote $F_I = |I| - S_I$.
- Consider now, for general $\bar{n}$, symmetric rules of the form

$$a(S_I, F_I),$$

Lemma (Implementable symmetric rules)

$\bar{a}(\cdot)$ is a reduced form publication rule that is implementable by such a symmetric rule iff it is of the form

$$\bar{a}(X) = \mathbf{1}(S \in \mathscr{S}),$$

where $\mathscr{S}$ is a union of intervals of length at least $\bar{n} - \bar{n}^{PC}$.

Optimal symmetric rules

- Minimal publication cutoff for the journal:

$$\underline{s}^{min}(n) = \min\left\{\underline{s}:\ E[\theta \big| S_{1,\dots,n} \geq \underline{s}] \geq \underline{\theta}\right\}.$$

Proposition (Optimal symmetric publication rule)

The optimal reduced-form publication rule that is symmetrically implementable takes the form

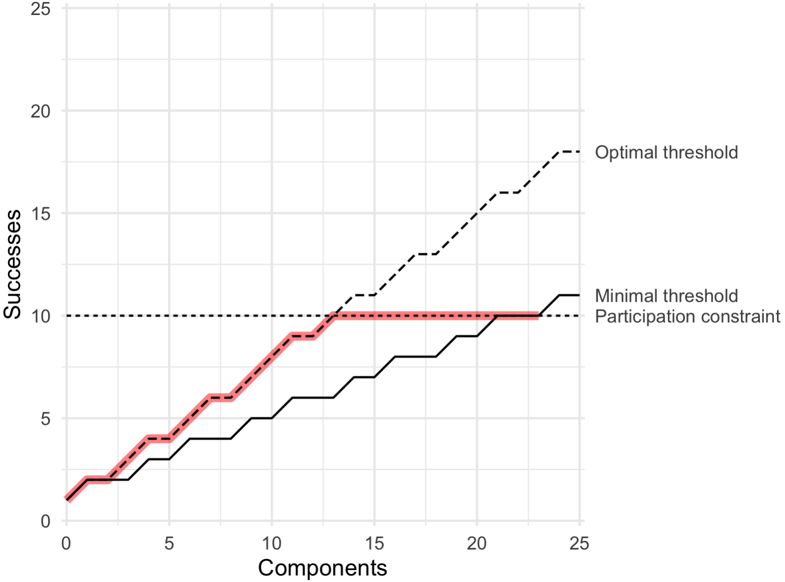
$$\bar{a} = \mathbf{1}(S \geq \min(\underline{s}^*, \bar{n}^{PC})),$$

if $\bar{n}^{PC} \geq \underline{s}^{min}$*, and can be implemented by*

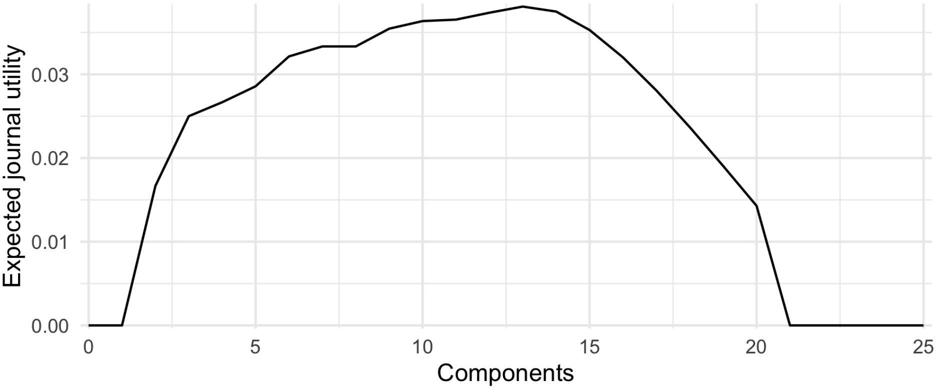
$$a = \mathbf{1}(S_I \geq \min(\underline{s}^*, \bar{n}^{PC})).$$

Otherwise the optimal publication rule is given by $a \equiv 0$*.*

Symmetric cutoff without PAP, uniform prior



Symmetric cutoff without PAP, uniform prior



PAPs and symmetric publication rules

The figure illustrates:

- If the number of components $\bar{n}$ is to the right of the maximum $\bar{n}^*$,
- then PAPs increase journal welfare
- by forcing the researcher to ignore all components $i > \bar{n}^*$.

General implementable rules

Theorem

The implementable publication functions $\bar{a}(x)$ *are exactly those that are of the form*

$$\bar{a}(x) = \mathbf{1}(x \in \cup_j C_{I_j, w_j}),$$

for some set of $\{(I_j, w_j)\}$, *where* $C_{I,w}$ *are the cylinder sets*

$$C_{I,w} = \{x : \; x_I = w\},$$

and $|I_j| = \bar{n}^{PC}$ *for all* j.

Optimal implementable rules (work in progress)

Conjecture (Optimal implementable publication functions)

Let $\bar{a}(x)$ *be optimal among implementable publication functions. Recall that* $\underline{s}^* = \min\left\{\underline{s} : \; E[\theta \big| S = \underline{s}] \geq \underline{\theta}\right\}$. *If* $\underline{s}^* \leq \bar{n}^{PC}$*, then* $\bar{a}(x) = 1(|x| \geq \underline{s}^*)$*. If* $\underline{s}^* > \bar{n}^{PC}$*, then*

$$\bar{a}(x) = 1(x \in \cup_{I \in \mathcal{I}} \{x' : \; x'_I = \mathbf{1}\})$$

for some family $\mathcal{I}$ *of index sets with with* $|I| = \bar{n}^{PC}$ *for all* $I \in \mathcal{I}$*. Furthermore,* $\mathcal{I}$ *is maximally spread out in the sense that if* $I_1 \neq I_2 \in \mathcal{I}$*, then all* $J \subseteq \{1, \ldots, n\}$ *with* $|J| = \bar{n}^{PC}$ *and* $|I \cap J| < |I_1 \cap I_2|$ *for all* $I \in \mathcal{I} \setminus \{J\}$ *are also in* $\mathcal{I}$*.*

Introduction

Baseline model

- Assumptions
- Implementability and optimality

Analysis

- A minimal example: $\bar{n} = 3$
- Symmetric publication rules
- General solution

Model variations

- Frequentist testing
- Multiple parameters / hypotheses
- Cost of observation

Conclusion

Model variation I: Frequentist testing

- Setup same as in the baseline model, except for the journal objective:
  - Consider the **null hypothesis** $\theta \leq \underline{\theta}$,
  - **significance level** $\underline{\theta}$.
  - $\Rightarrow X_i$ is a valid test.
- **First best** rule (uniformly most powerful test): Critical value $\underline{s}^{test}(\bar{n})$,

$$\bar{a}(X) = \mathbf{1}(S \geq \underline{s}^{test}(\bar{n})).$$

- When $\underline{s}^{test}(\bar{n}) > \bar{n}^{PC}$, the first best is **not implementable**.
- Second best: Use PAP to restrict $\bar{n}$ to the largest value such that $\bar{a}(X)$ is implementable.
- Our previous analysis carries over almost verbatim!

Model variation II: Multiple parameters / hypotheses

- Setup same as in the baseline model, except for the journal objective:

$$u^{jour}(a) = a \cdot \sum_{i \in I} (\theta_i - \underline{\theta}),$$

where there are parameters θ_i for every i.

- Joint distribution of data and parameters:

$$\begin{aligned} X_i | \theta_1, \ldots \theta_{\bar{n}}, \bar{\theta} &\sim Ber(\theta_i) \\ \theta_i | \bar{\theta} &\sim Beta(m\bar{\theta}, m(1-\bar{\theta})) \\ \bar{\theta} &\sim \pi, \end{aligned}$$

- Selective reporting distorts inference.
  - For large $\bar{n}$ or c, the first best is not implementable,
  - but a PAP allows to implement the second best.

Model variation III: Cost of observation

- Setup same as in the baseline model, except for the researcher objective:

$$u^{res} = a - c \cdot \bar{n}.$$

  - $\bar{n}$ is an additional choice parameter of the researcher.
  - c is private information of the researcher.
    (Now a cost of observation, not communication.)

- Now $\bar{n}$ is endogenous to $\underline{s}$, and uncertain for the journal.
  - For low realizations of c, the journal will over-publish.
  - Uncertainty over $\bar{n}$ hurts the journal.

- Using PAPs to reduce uncertainty over $\bar{n}$
  can again allow the journal to implement the second best.

Introduction

Baseline model
- Assumptions
- Implementability and optimality

Analysis
- A minimal example: $\bar{n} = 3$
- Symmetric publication rules
- General solution

Model variations
- Frequentist testing
- Multiple parameters / hypotheses
- Cost of observation

Conclusion

Summary

- Single agent (statistical) decision theory can not rationalize PAPs.
- Mechanism design allows us to study implementable statistical decision rules.
- In our model, PAPs are optimal when
  1. there are many researcher degrees of freedom
  2. and communication costs are high.
- Extensions of the baseline model:
  1. Researcher **private information** about signal validity.
  2. Replacing the journal objective by **size and power** of a statistical test.
  3. Alternative cost structure 1:
     The **journal bears** the communication **cost**.
  4. Alternative cost structure 2:
     The **researcher** bears a **cost for observing**, rather than reporting, components.
  5. **Multiple** parameters or **hypotheses**.

Thank you!