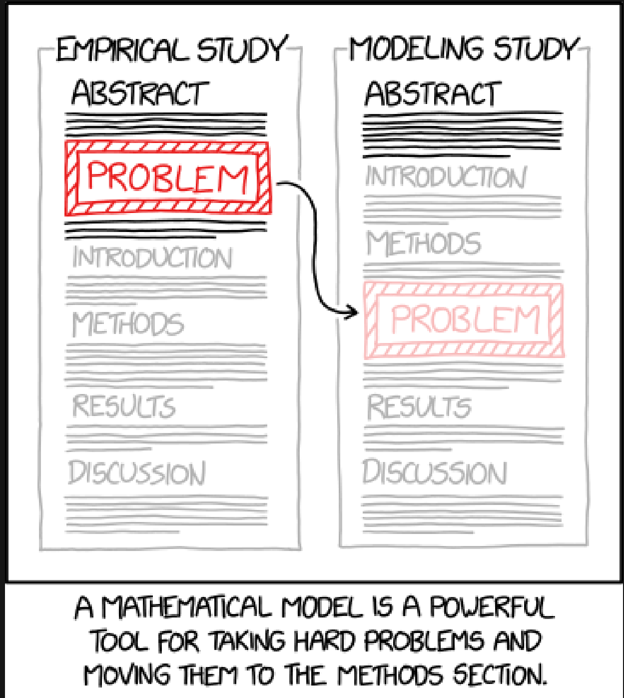
Comparing models of irrational behaviour against individual variance

Lenard Dome and Andy Wills

Apr 12, 2022

XKCD:2323



Structure

1. The framework
2. The tools
3. The phenomenon
4. The method
5. The findings

The framework: *global* model behaviour

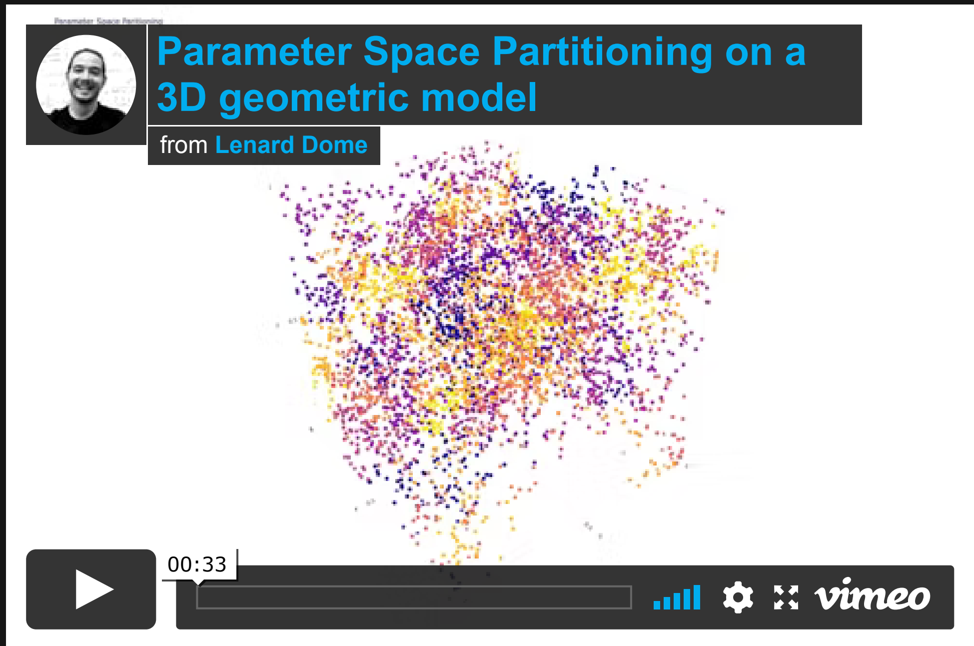
Roberts and Pashler (2000) pointed out three distinct areas not covered by a goodness-of-fit:

1. prediction What does the model say will happen?
2. heterogeneity Between-subject variability is not explained by goodness-of-fit.
3. *a priori* likelihood How likely that the model will be a good fit?

We evaluate models on these neglected areas of formal modelling.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. Psychological review, 107(2), 358.

The tools: parameter space partitioning



- *To model behavior, we need to know how models behave* (Pitt, Kim, MNavarro, Myung, 2006)
- MCMC method to define disjointed regions in the parameter space
- Tells us what unobserved results the model predicts

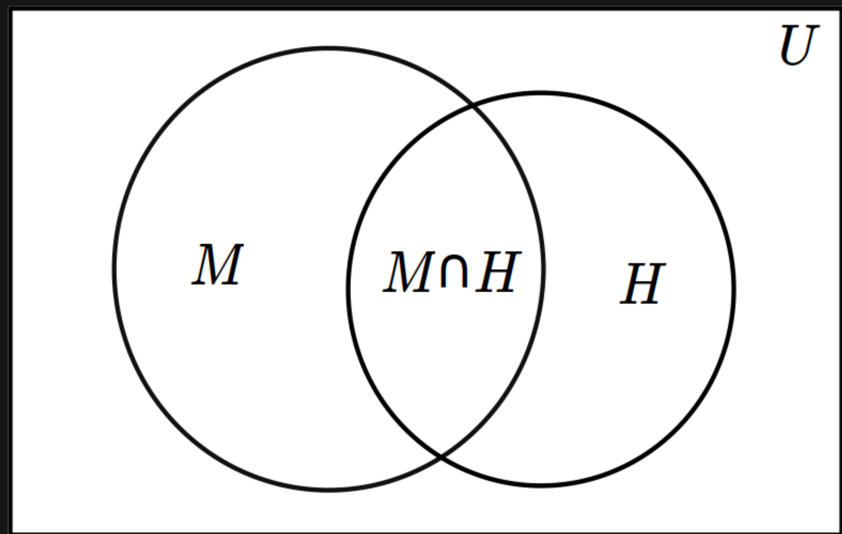
We implemented the technique in an R package: psp

R-CMD-check passing | downloads 3306 | cran v0.1 | license GPL (>= 3)

Buy me a coffee

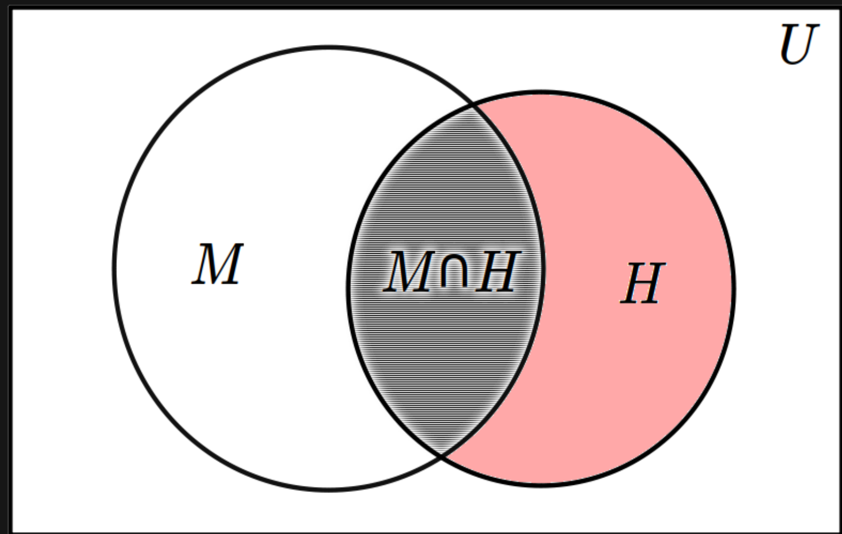
Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. Psychological Review, 113(1), 57.

The tools: g-distance



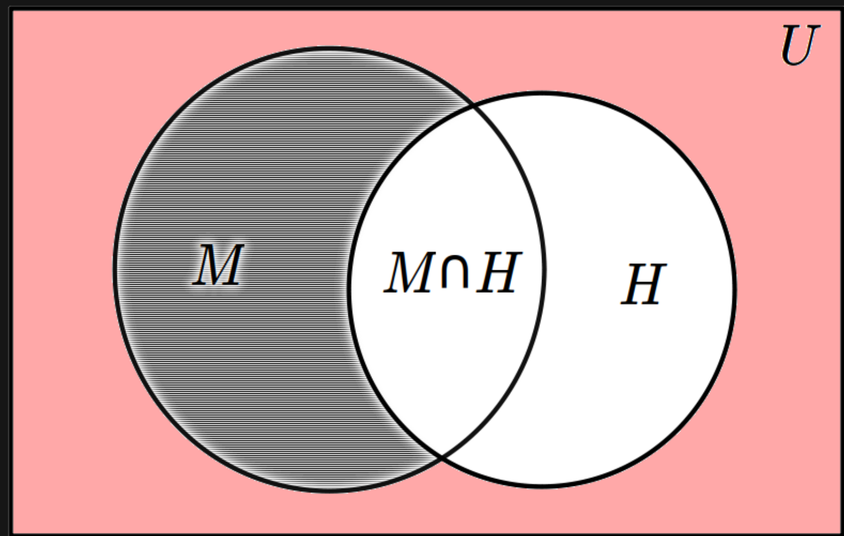
The relationship between these sets will be our measure

The tools: g-distance

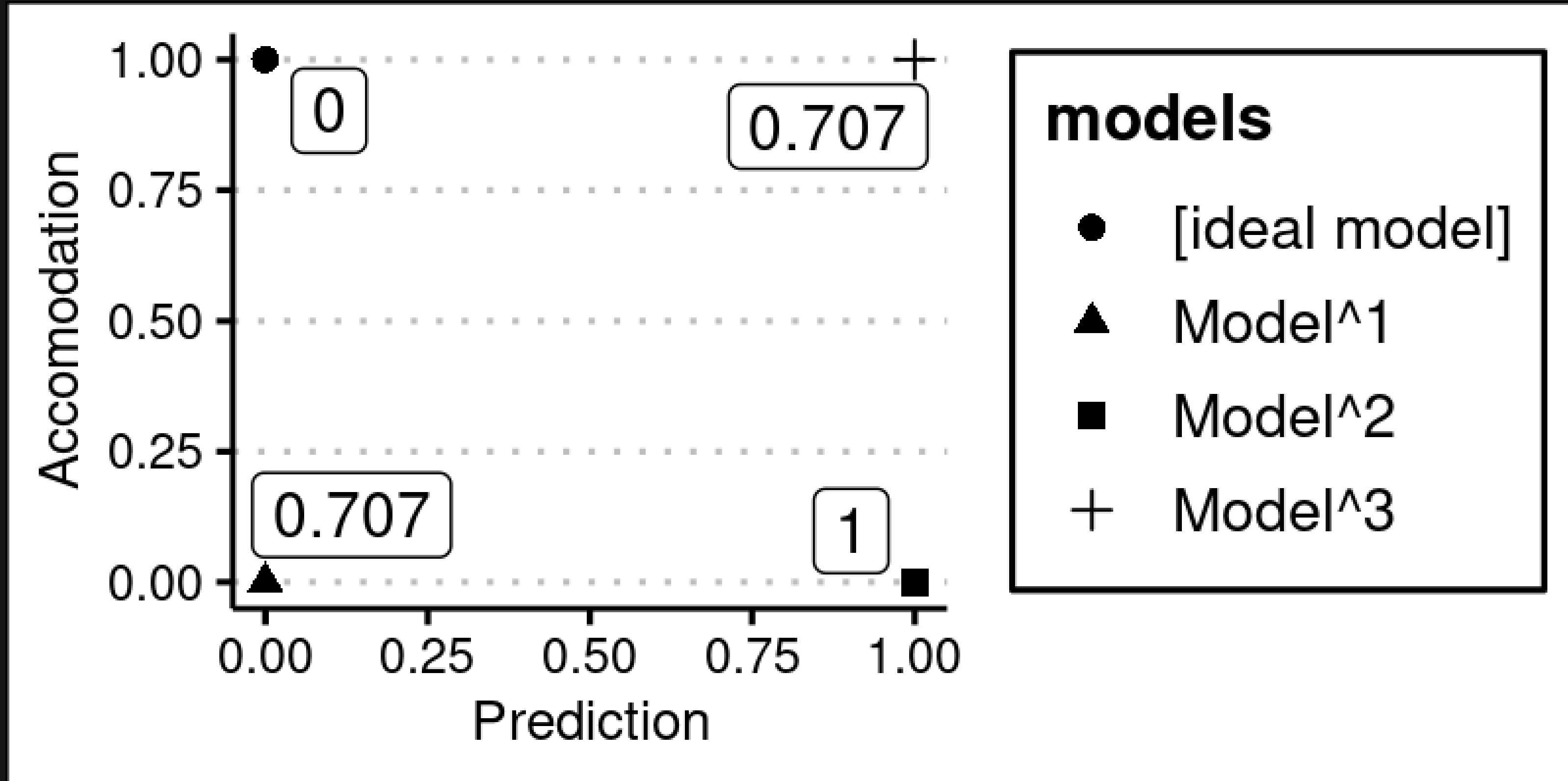


Sufficiency of **accomodation** $\alpha = \frac{|M \cap H|}{|H|}$

The tools: g-distance



Breadth of **prediction** $\beta = \frac{|M \cap H'|}{|H'|}$



---

g-distance is then the distance between our model and an [ideal model in an ideal scenario]

$$g = \sqrt{w_\alpha(1-\alpha)^2 + (1-w_\alpha)(0-\beta)^2},$$

The phenomenon: the inverse base-rate effect

Abstract Design

Training	Test
$AB \rightarrow common$ x3	A, B, C
$AC \rightarrow rare$ x1	BC

- Participants overwhelmingly prefer $BC \rightarrow rare$
- $BC \rightarrow rare$ is observed even when:
  - participants have high accuracy
  - participants have low accuracy
  - stimuli are pictures, colours, symptoms, person characteristics

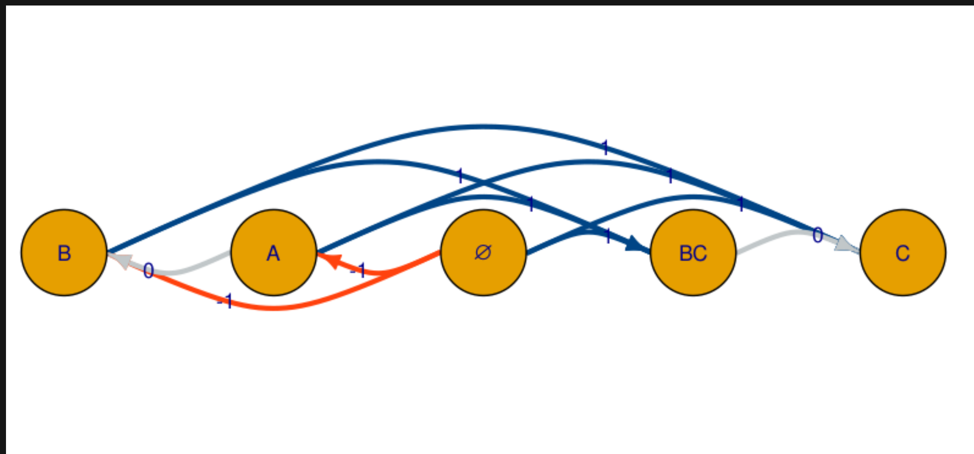
See Don, Worthy and Livesey (2021)'s excellent paper for a more thorough review.

Don, H. J., Worthy, D. A., & Livesey, E. J. (2021). Hearing hooves, thinking zebras: A review of the inverse base-rate effect. Psychonomic Bulletin & Review, 28(4), 1142-1163.

Method: ordinal patterns

Ordinal patterns are a qualitative description of a relationship between an independent variable as defined by at least one of the dependent variables.

Directed Graph



Inequality Matrices

	∅	A	B	BC	C
∅		1	1	1	-1
A			0	-1	0
B				-1	0
BC					0
C					

Method: procedure

Empirical

- We implemented IBRE in the simplest way possible
- Stopped data collection at 117 participants
- Long test phase
- Confirmed group-level resultsfo $BC \rightarrow rare$ > $BC \rightarrow common$

Computational

- All models were implemented in catlearn
- Ran PSP for each participant on different ordinal complexities
- Models we tested: EXIT, NNCAG, DGCM

The findings: results

> The computer was on a
> figurative 🔥

- There are 3 possible patterns.
- Hmans showed all three.

```
## [[1]]
##     ∅ BC
## ∅  NA -1
## BC NA NA
##
## [[2]]
##     ∅ BC
## ∅  NA  1
## BC NA NA
##
## [[3]]
##     ∅ BC
## ∅  NA  0
## BC NA NA
```

	pattern	*g-distance*	accomodation	prediction
EXIT	BC	0.47	1.00	0.67
NNCAG	BC	0.43	0.91	0.55
DGCM	BC	0.47	1.00	0.67

The findings: results (increasing complexity)

- There are 19 possible patterns
- Humans showed 17 of them.

```
## [[1]]
##     ∅  A BC
## ∅  NA  1  1
## A  NA NA -1
## BC NA NA NA
```

	pattern	*g-distance*	accomodation	prediction
EXIT	BC, A	0.67	0.20	0.41
NNCAG	BC, A	0.61	0.18	0.19
DGCM	BC, A	0.53	0.38	0.42

The findings: results (increasing complexity more)

- There are 171 possible patterns
- Humans showed 41 of them.
- EXIT showed 70 of them (41%) - independent of trial order.

```
## [[1]]
##     ∅  B BC  C
## ∅  NA  1  1  0
## B  NA NA  0 -1
## BC NA NA NA -1
## C  NA NA NA NA
```

	pattern	*g-distance*	accomodation	prediction
EXIT	BC, B, C	0.67	0.08	0.14
NNCAG	BC, B, C	0.65	0.09	0.07
DGCM	BC, B, C	0.57	0.20	0.12

The findings: conclusion

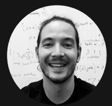
Methodological

- Developed a novel method to evaluate aspects of computational models often neglected

Theoretical

- Provided evidence against EXIT, the best model of IBRE
  - *EXIT has an issue of flexibility when trial-order is not accounted for*
- The most adequate model is a non-associative learning model, DGCM
  - *but it suffers from a metatheoretical issue*

The End



Lenard Dome

Buy me a coffee

- lenard.dome@plymouth.ac.uk
- @lenarddome
- @lenarddome

Andy Wills

- andy.wills@plymouth.ac.uk
- @ajwills72
- @ajwills72

UNIVERSITY OF PLYMOUTH
School of Psychology

UKRI Economic and Social Research Council

BRIC
BRAIN RESEARCH & IMAGING CENTRE