Hypothesis Testing with Computational Models

Hang Xiong — ETH Zurich

Georgiy Bobashev — RTI International

Outline

- Hypothesis in computational models
- The procedure of hypothesis testing we propose
- An application of the procedure
- Discussions

Hypothesis in computation models

	Statistical model	Computational model (e.g. ABM)
Components	$y = f_{math}(x)$	$output = f_{comp}(agents, rules)$
Hypothesis	Influence of *x* on *y*	Influence of a *rule* on *output*
Test	Correlation btw x and y	Prob of the parameter associated with the rule

Computation models are built on hypothesized a rule of how agents act e.g. peer effects

The rule often can be indicated by a parameter (or a set of parameters) e.g. the threshold value in Schelling's segregation model

Hypothesis in computation models (cont'd)

Suppose a computational model

$$y = F(x, \alpha, \beta)$$

- y - outputs
- x - inputs
- α - the parameter indicating the hypothesized effect
- β - the known parameters in addition to α
- $\theta = \{\alpha, \beta\}$

Test if parameter α is significant

Hypothesis testing approaches

Frequentist approach

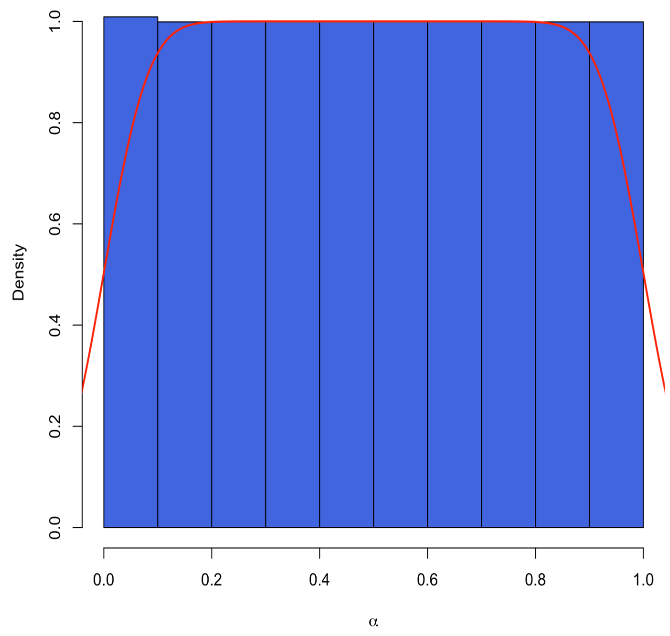
- Given the H_0, what is the probability of more extreme data?
- Some sort of Goodness of Fit (GOF) test to either accept or reject H_0

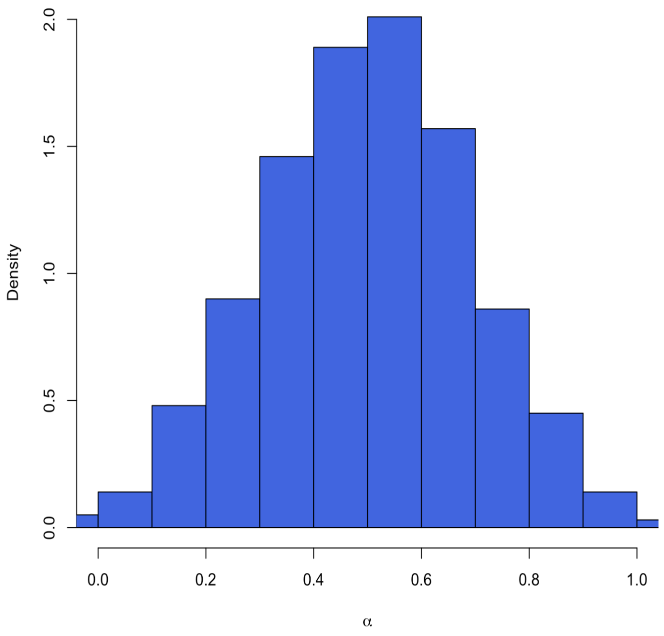
Bayesian approach

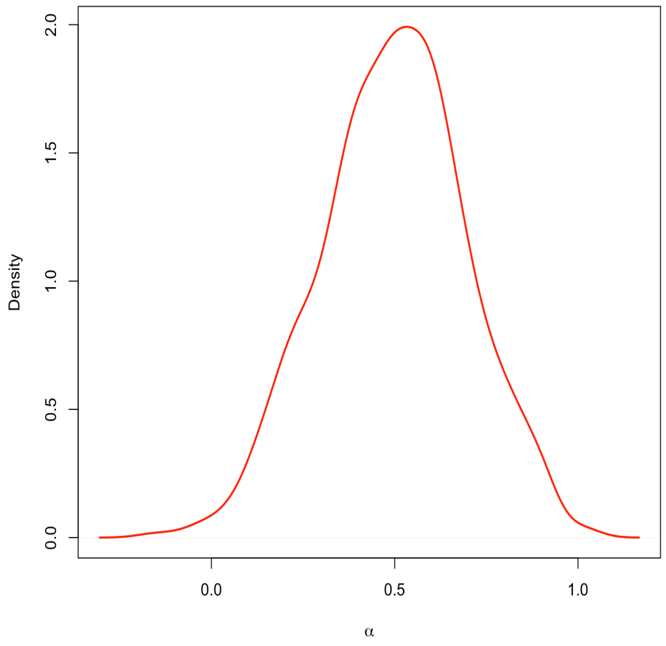
- Given the data, what is the probability of H_0?
- Estimate the posterior distribution of the parameter and evaluate where the H_0 is in that distribution

General procedure of Bayesian approach

Assume prior distribution of α ⇒ Estimate α from the best fit to the data ⇒ Compute the distribution of α using bootstrap







Steps of hypothesis testing (Bayesian approach)

1. Construct the null/alternative hypothesis
2. Define the criterion function for estimating the parameters
3. Assume the prior parameter distribution
4. Estimate the posterior parameter distribution
5. Test the probability of the null hypothesis

Construct the null/alternative hypothesis

Null hypothesis: the probability the effect is too small is higher than a threshold

e.g., $P(\alpha < 0.05) > 10\%$

Alternative hypothesis: the probability the effect does not exist is lower than a probability

e.g. $P(\alpha < 0.05) \leq 10\%$

Define the criterion function

For determining the optimal parameter combination that minimize the discrepancy between the simulated results and the observed data.

A simple form of the function:

$$agrmin_{\theta \in \Theta}(y_{sim}(\theta) - y_{obs})^2$$

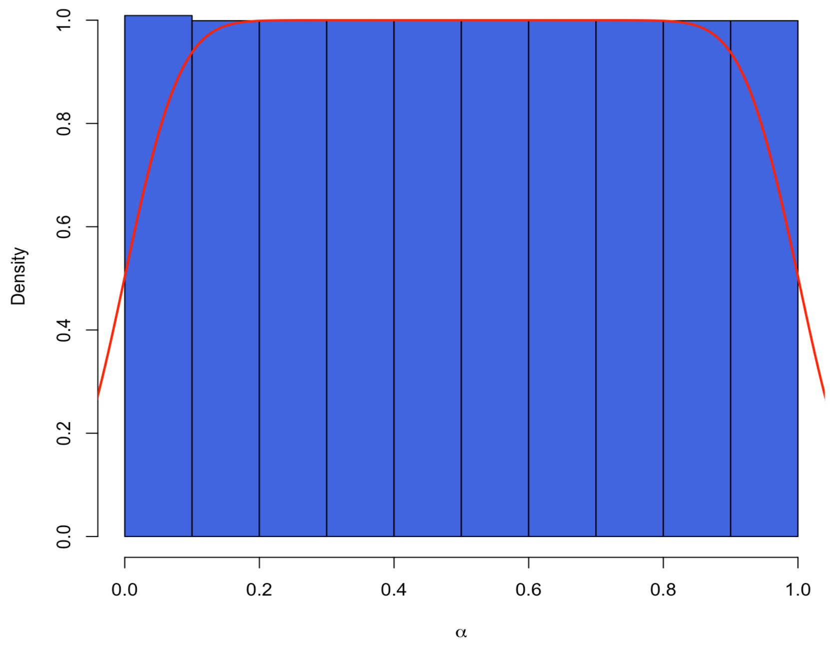
Method of simulated moments can be used to derive the features of sample statistics as a basis for comparing to the data.

Assume the prior parameter distribution

Specify the parameter space

Define the prior distribution for α, as well as other parameters, based on prior knowledge

In case of no prior knowledge, such as our case, assume non-informative distributions i.e. all parameters uniformly distributed in the parameter space



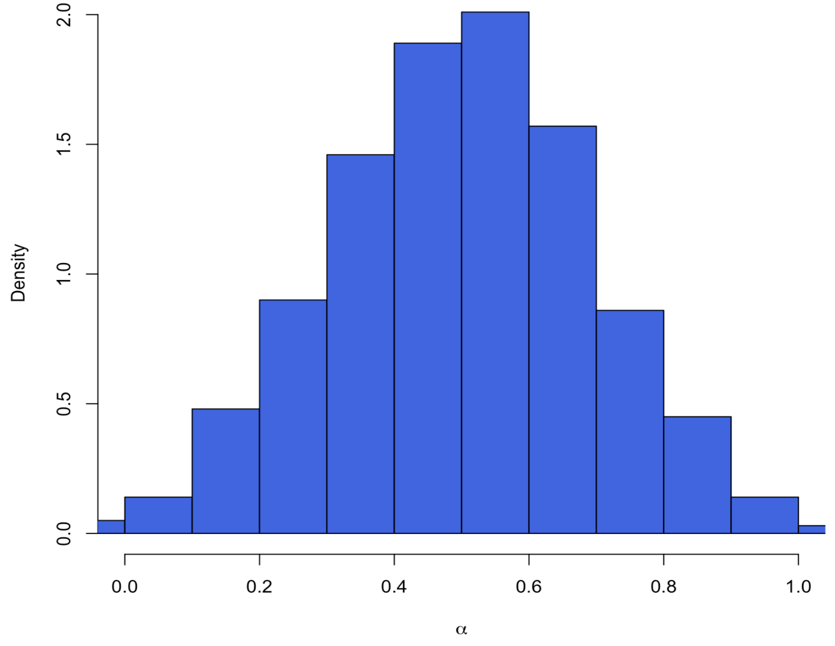
Use Goodness of Fit (e.g. ML or MSE) to Estimate the parameter value

Find the optimal parameter combination for each set of data

$$\alpha_i^* = agrmin_{\theta \in \Theta}\left(y_{sim,i}(\theta_i^*) - y_{obs,i}\right)^2$$

Use optimal parameter search algorithms e.g. exhaustive grid search.

Narrow down and refine the search space.

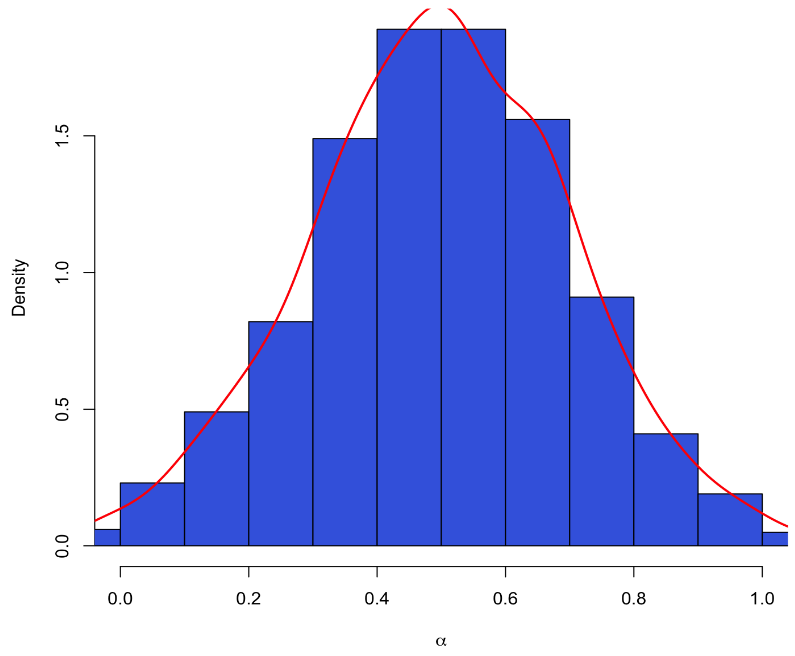


ETH zürich RTI INTERNATIONAL

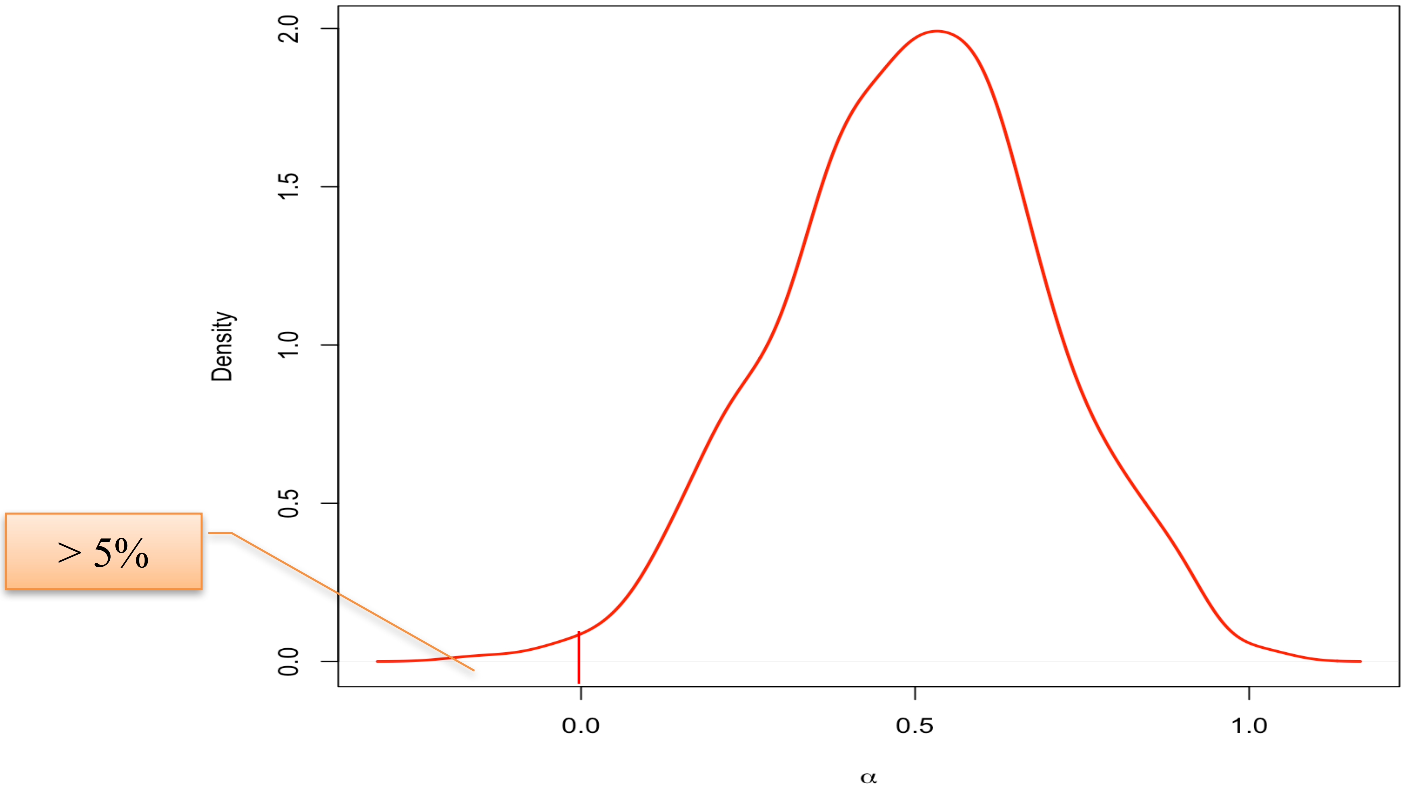
Estimate the posterior parameter distribution (cont’d)

Require large sample size

In the case of small sample size, use techniques such as bootstrap

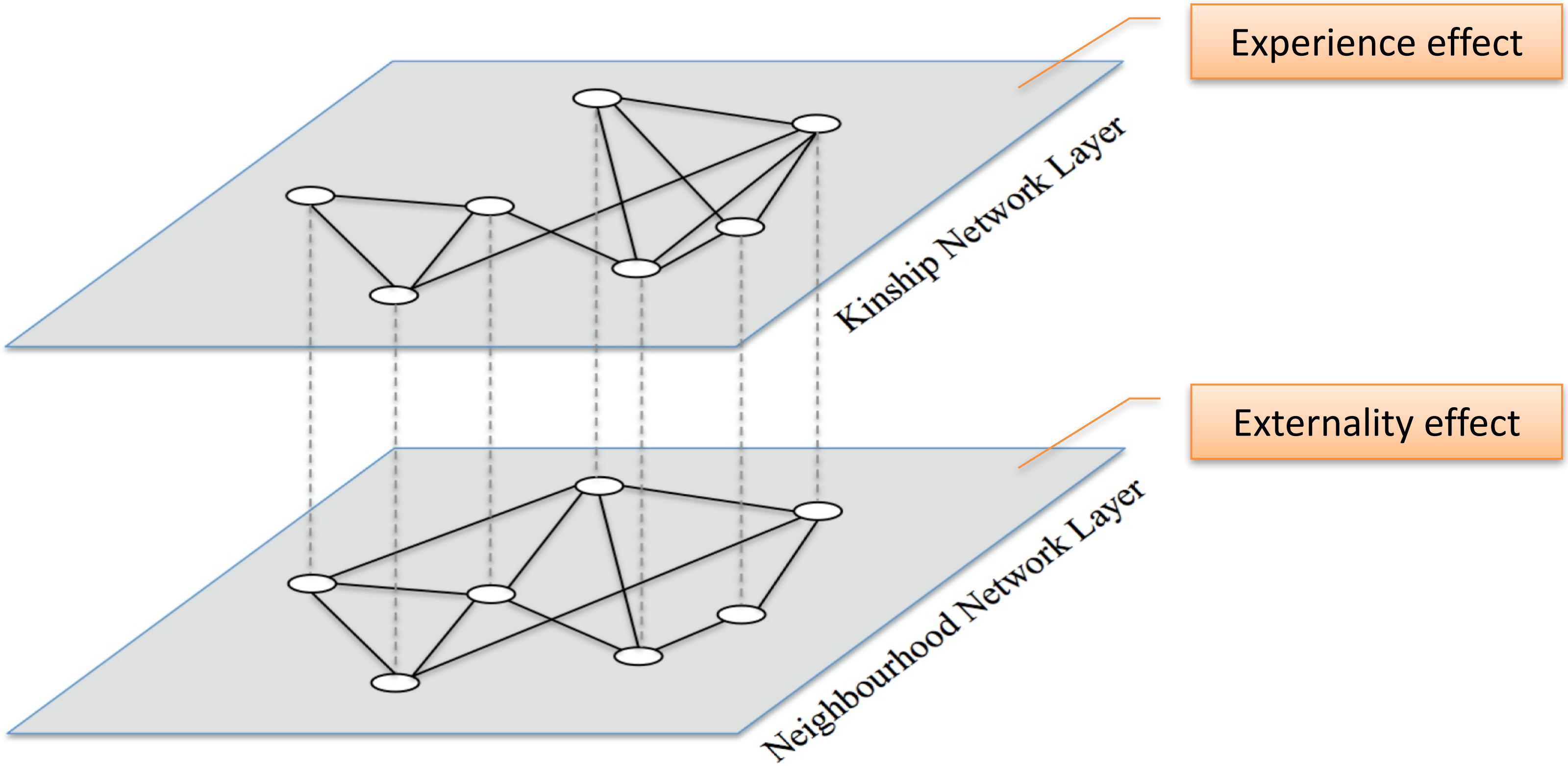


Test the probability of the null hypothesis



An application: peer effects in the diffusion of new crops

Two hypothesized peer effects



ETH zürich RTI INTERNATIONAL

The two effects are each indicated by a parameter

- First, only ***experience effect*** is modelled. A household's probability to adopt is given by

$$\log\left(\frac{p_{it}}{1-p_{it}}\right) = \alpha + X_r'\beta + \gamma F_{it}$$

where $F_{it} = \dfrac{\#\ \textit{of adopting relatives}}{\#\ \textit{of relatives}}$ represents the experience effect, and γ is *experience coefficient*.

- Then, ***externality effect*** is included. A non-adopting household adopts when

$$H_{it} > h$$

where H_{it} is the fraction of adopting land plot neighbours, and h is *externality threshold*.

ETH zürich RTI INTERNATIONAL

Null hypothesis and criterion function

- **Step 1**

  null hypothesis: $P(\gamma = 0) > 5\%$; $P(h = 0) > 5\%$

- **Step 2**

  Estimate parameters by minimising a criterion function that is a quadratic form of the distance between empirical adoption rate in each year for each village (i.e., moments) and the adoption rate predicted by the model for a specific combination of parameters, namely

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{r=1}^{R} \sum_{t=1}^{T} |\overline{Y}_{t,r}(\theta) - Z_{t,r}|$$

where $\overline{Y}_{t,r}(\theta) = \frac{\sum_{s=1}^{S} Y_{t,r}(s,\theta)}{S}$

Prior distributions of parameters

- **Step 3**

1. Experience Model

   (i) Pure Experience Effect:
   $\alpha \in [(-3.00, -2.45, 0.05), (-2.50, -2.00, 0.02)]$
   $\beta \in [(0.10, 0.60, 0.01), (0.70, 0.90, 0.02)]$
   $\gamma \in [(4.00, 4.50, 0.02), (4.55, 5.00, 0.05)]$

   (ii) Close Relatives:
   $\alpha \in [(-2.50, -2.05, 0.05), (-2.00, -1.70, 0.02), (-1.65, -1.50, 0.05)]$
   $\beta \in [(0.10, 0.35, 0.05), (0.40, 1.00, 0.02)]$
   $\gamma \in [(3.00, 3.15, 0.05), (3.20, 3.40, 0.02), (3.60, 4.00, 0.05)]$

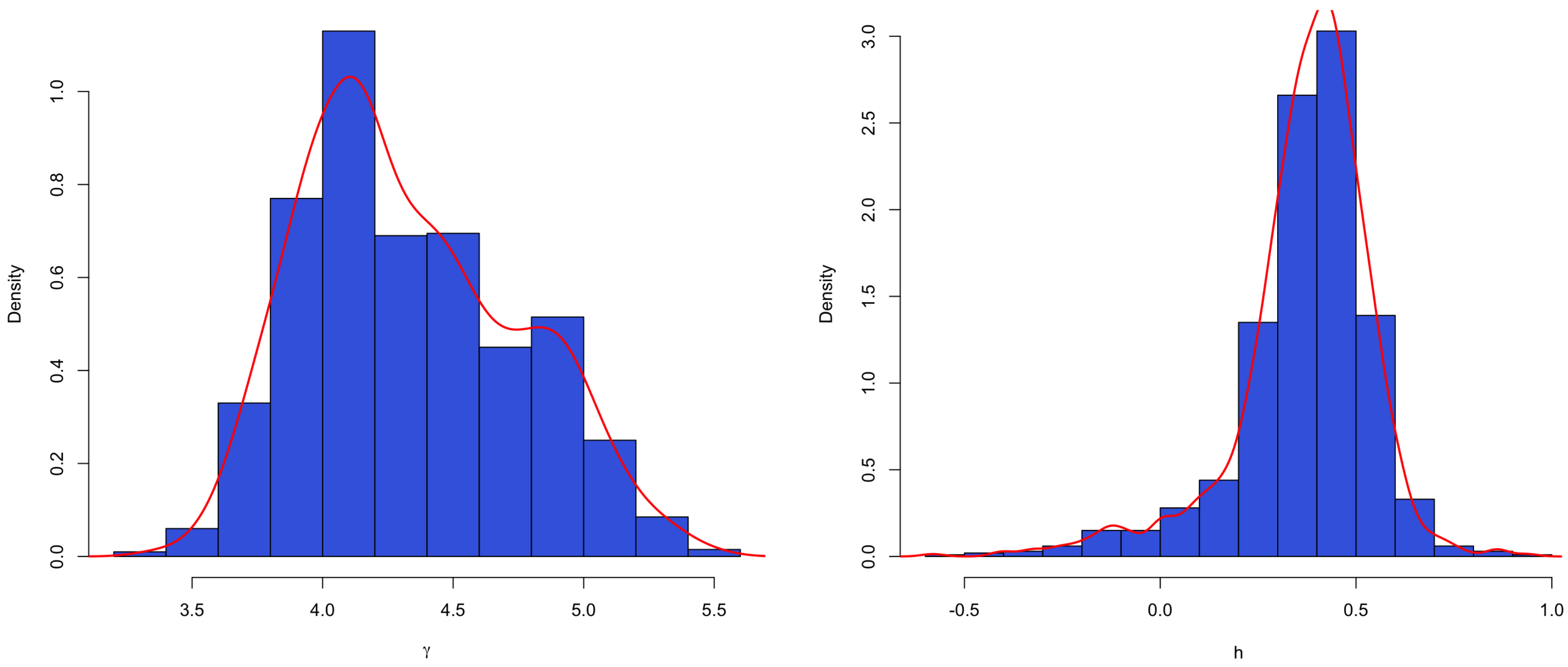
   (iii) Proximity in Age:
   $\alpha \in [(-2.50, -2.30, 0.10), (-2.20, -1.80, 0.05), (-1.70, -1.50, 0.10)]$
   $\beta \in [-0.25, 0.25, 0.05]$
   $\gamma \in [4.00, 5.00, 0.05]$
   $\eta \in [(0.025, 0.50, 0.025), (0.55, 1.00, 0.05)]$

2. Experience-Externality Model

   (i) All Years:
   $\alpha \in [-3.00, -2.00, 0.05]$
   $\beta \in [-0.50, 1.50, 0.05]$
   $\gamma \in [3.00, 4.00, 0.05]$
   $\delta \in [(0.02, 0.50, 0.02), (0.55, 1.00, 0.05)]$

Posterior distributions of parameters

- **Step 4**



Test the probability of the null hypotheses

- **Step 5**

	γ (1)	η (2)	h (3)
Panel A: Experience Model			
Pure Experience Effect	4.10		
Standard Error	[0.2569]		
99% CI of Bootstrap Distribution	[3.84, 5.16]		
Close Relatives	3.32		
Standard Error	0.2425		
99% CI of Bootstrap Distribution	[2.73, 3.98]		
Proximity in Age	4.50	0.275	
Standard Error	[0.2964]	[0.1231]	
99% CI of Bootstrap Distribution	[3.72, 5.25]	[−0.04, 0.60]†	
Panel B: Experience-Externality Model			
All Years	3.60		0.30
Standard Error	[0.2929]		[0.2226]
99% CI of Bootstrap Distribution	[2.83, 4.34]		[−0.20, 0.95]
Frist 4 Years	4.80		0.42
Standard Error	[0.2979]		[0.2339]
99% CI of Bootstrap Distribution	[3.76, 5.30]		[−0.13, 1.08]
Last 4 years	3.10		0.48
Standard Error	[0.4477]		[0.0788]
99% CI of Bootstrap Distribution	[2.79, 4.88]		[0.13, 0.50]

† 95% CI of Bootstrap Distribution is [0.04, 0.52].

Discussions

This procedure is a collection of often-implemented exercises, but it aims to formulate a practice of statistical inference with computational models.

It can be viewed as an approach of model validation with real data.

It requires a large sample size, but techniques such as bootstrap can help.

It is especially useful for testing the effects that statistical models cannot well capture e.g. the interactive effects.

Thank you for your attention. Comments, Questions, Suggestions?

hxiong@ethz.ch