Inferring social learning processes from empirical data: Which details matter?

Anne Kandler

June 10, 2024

Department of Human Behavior, Ecology, and Culture
Max Planck Institute for Evolutionary Anthropology

TI(CE) LAB
Theory in Cultural Evolution

Introduction

Introduction

The behavioural adaptations that explain the success of our species are widely thought to be cultural: they are transmitted among individuals by **social learning** and have **accumulated over generations**.

Introduction

This demographic success is driven mainly by cultural and not genetic evolution.

There are significant patterns of geographical variation in human behaviour at the social groups level: neither genes, nor environment, nor a combination of the two is sufficient to explain them.

To understand our demographic success (and human behaviour in general), we must understand the process of cultural evolution.

→ The core idea of cultural evolution is that cultural change constitutes an evolutionary process that shares fundamental similarities with — but also differs in key ways from — genetic evolution.

Introduction

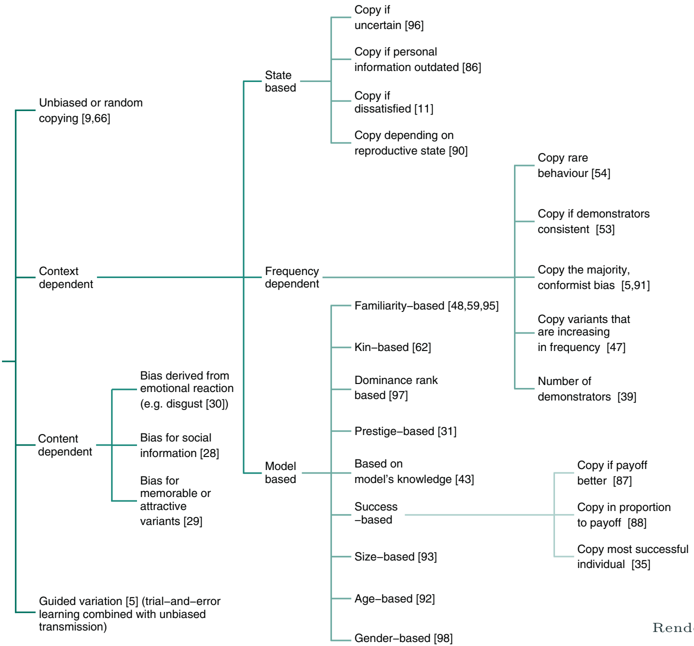
The behavioural adaptations that explain the success of our species are widely thought to be cultural: they are transmitted among individuals by **social learning** and have **accumulated over generations**.

Social learning/cultural transmission is defined as learning through observations of, or interactions with other individuals of the population or its products Heyes (1994), Hoppitt and Laland (2013)

Introduction

→ BUT there is no single unique way on how to learn socially e.g. Laland (2004)!

Learning/transmission biases



Rendell et al. (2011)

Introduction

→ Understanding how (and when) social learning is used is one of the central challenges in cultural evolution

Previous work

Research to establish the presence of particular learning strategies in human populations has mainly centered around

- **Laboratory-based approaches**: 'Microsocieties' (e.g. Coultas 2004, Baum et al. 2004, McElreath et al. 2008, Morgan et al. 2012) and diffusion chain experiments (e.g. Mesoudi and O'Brien 2008, Caldwell and Millen 2008, Kirby et al. 2008)
- **Inference-based approaches:** adoption curves (e.g. Cavalli-Sforza and Feldman 1981, Boyd and Richerson 1985, Rogers 2003, Henrich 2001, Reader 2004, Kendal et al. 2007, Hoppitt et al. 2010), power-law distributions (e.g. Hahn and Bentley 2003, Herzog et al. 2004, Bentley et al. 2004, Mesoudi and Lycett 2009), model selection frameworks (McElreath et al. 2008, Franz and Nunn 2009, Hoppitt et al. 2010)
- **Modelling-based approaches** (e.g. Feldman, Cavalli-Sforza et al., Boyd, Richerson et al.)

Previous work

Modelling-based approaches (e.g. Feldman, Cavalli-Sforza et al., Boyd, Richerson et al.)

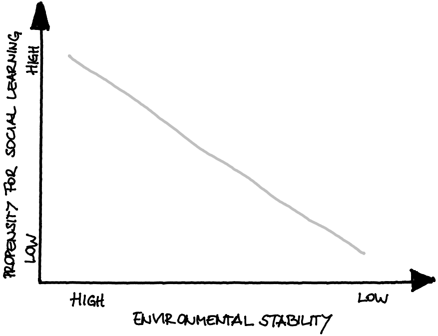
→ formal models of cultural evolution, along the lines of those models that stimulated the modern synthesis in biology, appeared in the 1970s and 1980s

→ applied the same bookkeeping methods developed by Fisher, Haldane, Wright and others to culture on the assumption that biological and cultural change are at heart both Darwinian systems

Previous work

Modelling-based approaches (e.g. Feldman, Cavalli-Sforza et al., Boyd, Richerson et al.)

→ Evolutionary advantage/disadvantage of social learning over its counterpart, asocial learning (innovation) in unstable environments



The usefulness of social learning decreases when environments become more and more unstable Aoki and Feldman (2014) for a review of the literature

Previous work

Modelling-based approaches (e.g. Boyd and Richerson et al., Feldman et al.)

→ Evolutionary advantage or disadvantage of social learning over its counterpart, asocial learning (or innovation)

→ Learning strategies that are expected in human populations especially in spatially and temporally changing environments at equilibrium

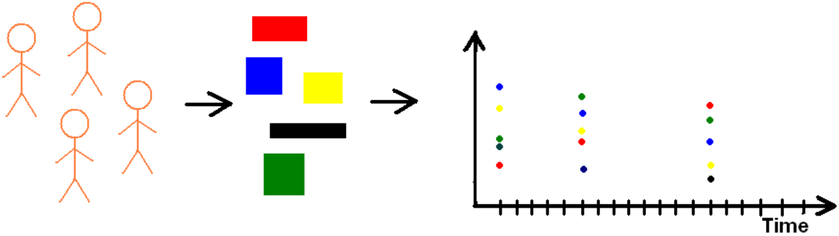
→ Important insight into what human populations are expected to do (at least based on the models used) if the cultural system is at equilibrium

But it often hard to verify these theoretical predictions empirically as fine-grained data detailing who is learning from whom and why is only rarely available.

→ We still don't know how humans learn socially (based on the available empirical data)

Available data (at least in archaeological and anthropological applications): population-level frequencies detailing the usage or occurrence of different variants of a cultural trait at one or different points in time

Pattern to process problem (inverse problem)



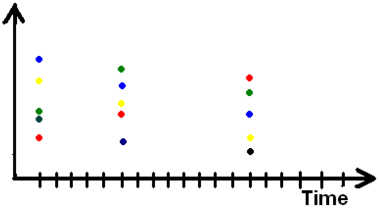
To be inferred: underlying processes of social learning

Observed: population-level frequencies of different variants of a cultural traits

Introduction

Generative inference framework

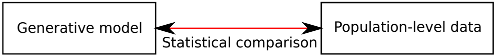
Generative model



To be inferred: underlying processes of social learning

Observed: population-level frequencies of different variants of a cultural traits

Generative inference framework



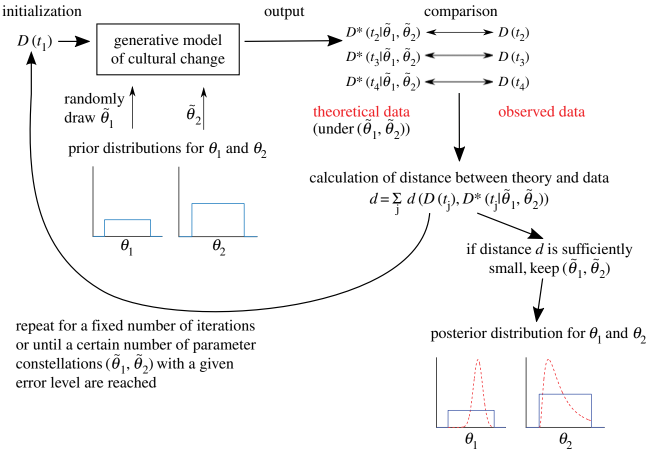
i. **Generative model:** (Non-equilibrium) framework capturing the main cultural and demographic dynamics of the cultural system

→ Frequencies of different cultural variants present in a population under the assumed learning hypothesis

→ Causal relationship between learning processes and observable frequency pattern

ii. **Statistical comparison:** Conclusions about which (mixtures of) learning processes are consistent with the observable frequency data and which are not.

Generative inference framework



Kandler and Powell (2018); Example of an application to archaeological data: Kandler and Shennan (2016)

Generative inference framework

→ How well the inverse problem can be solved depends on both, the quality of the empirical data and the "appropriateness" of the generative model

Generative inference framework

Precise inference results can only be achieved if the generative model is indeed a meaningful reflection of the cultural system!

→ Move away from simply adopting models of biological evolution towards developing a bespoke theory of cultural evolution that is applicable to the available data

→ BUT we cannot keep simply increasing modelling complexity

→ "What" needs to be included in the model?

Model (mis-)specification problem

Most models of cultural evolution have been adopted from biological evolution and assume:

- no age structure of the population,
- no social structure (i.e. no restriction of interactions between individuals through the interaction network),
- no learning of packages of traits,
- equilibrium,
- **individuals possess no memory** (i.e. if individuals learn a new cultural variant they immediately forget the one previously held),
- $\cdots$?

Model (mis-)specification problem

Most models of cultural evolution have been adopted from biological evolution and assume:

- no age structure of the population Kandler, Fogarty, Karsdorp (2023),
- no social structure (i.e. interactions between individuals are controlled by an interaction network) Smolla et al. (in prep),
- no learning of packages of traits Yeh et al. (2019),
- equilibrium Kandler, Crema (2018),
- **individuals possess no memory** (i.e. if individuals learn a new cultural variant they immediately forget the one previously held),
- $\cdots$?

$\rightarrow$ Which demographic and/or cultural processes have a sizeable impact on the dynamic of cultural change and therefore have to be included into models of cultural evolution?

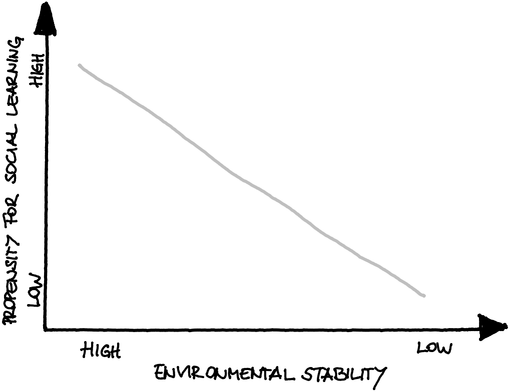
Model (mis-)specification problem

Most models assume cultural traits to be similar to genetic traits, i.e. each individual can adopt only one variant of the cultural trait at any point in time, there is no "memory" included in the models. But humans are clearly different!

$\rightarrow$ How does this modelling assumption influence e.g. results about the importance of social learning in changing environments?

Model (mis-)specification problem

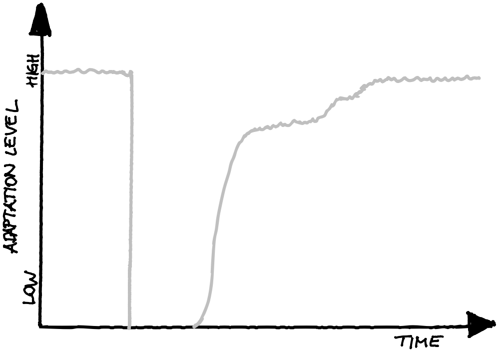
Answer from theoretical cultural evolutionary research:
through a combination of asocial (innovation) and social learning



The usefulness of social learning decreases when environments become more and more unstable Aoki and Feldman (2014) for a review of the literature

Model (mis-)specification problem

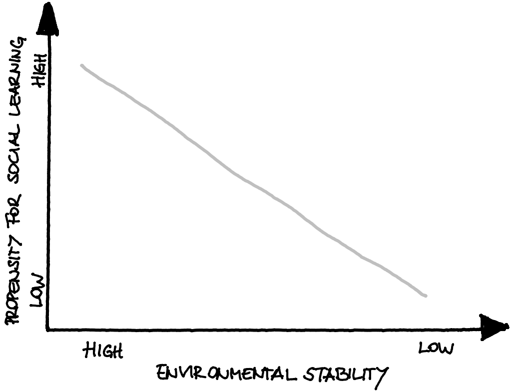
High environmental stability



Adaptation through *de novo* innovation

Model (mis-)specification problem

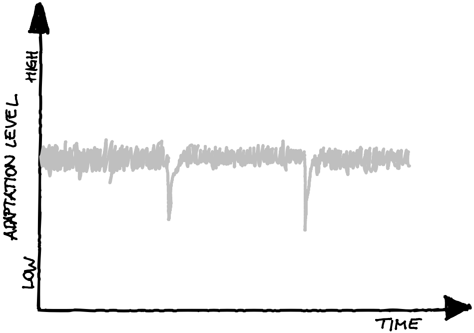
Classical result:



The usefulness of social learning decreases when environments become more and more unstable Aoki and Feldman (2014) for a review of the literature

Model (mis-)specification problem

Low environmental stability



Adaptation through standing variation

Memory

Human cognitive capacity allows for memory processes, i.e. the recall of past experiences — in our case cultural variants and their adaptation level in experienced environments

→ expressed and unexpressed cultural traits

Can “memory” contribute to the efficiency of the cultural adaptation process?

Ammar, Fogarty and Kandler (2023)

Model

Environment: can assume 2 distinct states $\mathbf{E} = \{1, 2\}$
$\rightarrow$ in each time step the environment can change with probability p_{env}

Cultural variants: each variant is only adapted to a single environmental state and possesses an adaptation value $a_i \in (0, 1)$

Individuals:

- possess a cultural memory comprised of their knowledge about a number of cultural variants and the adaptation level they provided in the experienced environmental states
- are characterised by their propensity, ξ_j, to engage in social learning and its propensity, φ_j, to forget knowledge

Model

Cultural dynamics:

- Choice of cultural variant to express
  - Evaluation of memory: each individual j goes through their memory and chooses a variant proportional to their adaptation level in the current environment (expected adaptation value: a_{memory})
    - With probability a_{memory} it expresses the chosen variant
  - Decision to learn (or not):
    - With probability $1 - a_{\text{memory}}$ individual j learns
      $\rightarrow$ it engages in social learning with probability ξ_j and in innovation with probability $1 - \xi_j$

      Social learning: payoff-biased learning

      Innovation: introduction of variant i with adaptation level $a_i \sim \mathcal{U}(0, 1)$ in the current environment

      $\rightarrow$ variant i is expressed and added to individual j's memory

Model

Cultural dynamics:

- Forgetting: each individual forgets a cultural variant, contained in its cultural memory, with probability φ_j

  $\rightarrow$ Variant to be forgotten is chosen inversely proportional to the number of times it has been expressed.

Model

Demographic dynamics: Moran model

$\rightarrow$ one individual is chosen for death and one for reproduction (death occurs at random, reproduction occurs proportional to the adaptation level of the expressed variant in the last time step)

Intergenerational transmission dynamic:

$\rightarrow$ individual j passes on its ξ and φ-values with probability $1 - \mu_\xi$ and $1 - \mu_\varphi$, respectively. Otherwise mutations occur

$$\xi_j + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma_\xi^2),$$

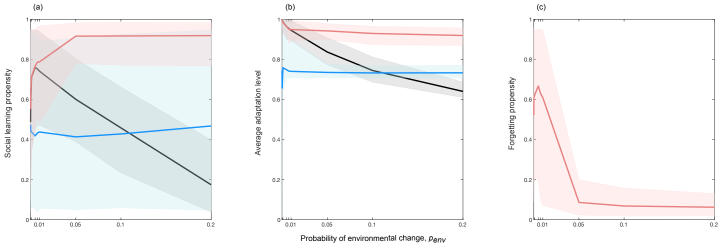
$$\phi_j + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma_\varphi^2).$$

$\rightarrow$ individual j passes on its memory

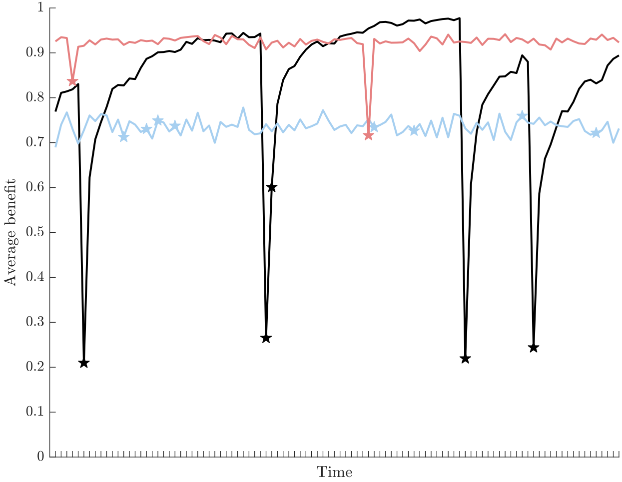
Influence of memory

$\varphi_j = 1, \ \forall j \rightarrow$ no memory, classical social learning result

Adaptation level $\qquad \xi \qquad \varphi$



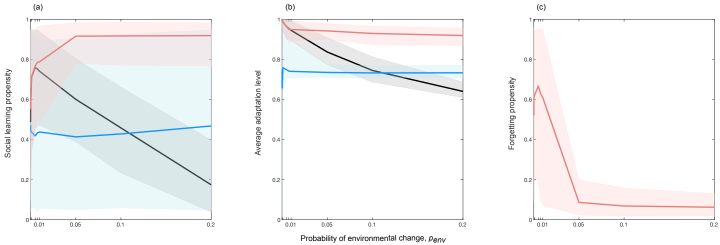
Influence of memory



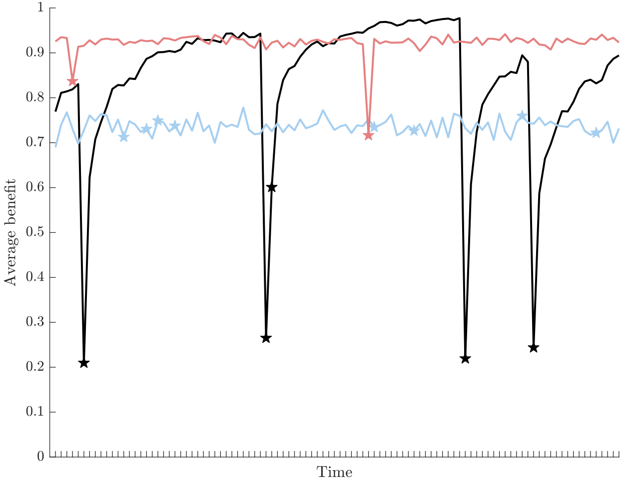
Influence of memory

$\varphi_j = 0, \ \forall j \to$ full memory

Adaptation level $\qquad\qquad \xi \qquad\qquad \varphi$



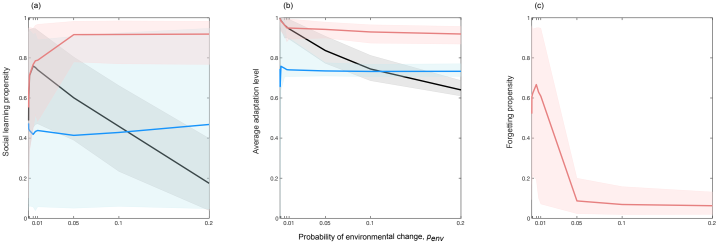
Influence of memory



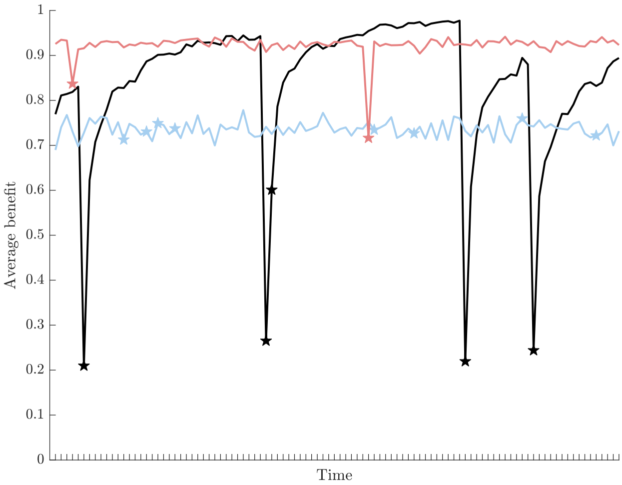
Influence of memory

Evolving memory

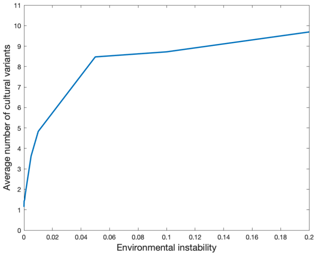
Adaptation level $\qquad$ ξ $\qquad$ φ



Influence of memory

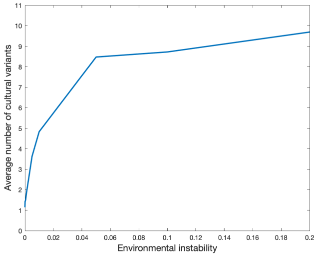


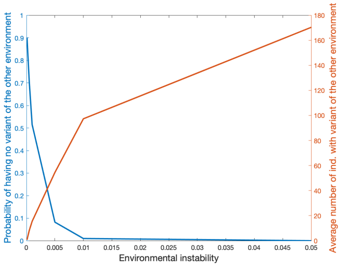
Structure of memory



→ Individual memories stay relatively small

Structure of memory





→ Individual memories stay relatively small

→ Not every individual contains information about the other environment in the memory

Conclusions

The inclusion of memory, i.e. individual collections of cultural variants in the past, can change the relationship between social learning and environmental uncertainty.

Conclusions

Demographic and cultural properties of the population can have a great influence on the dynamic of cultural change.

→ We need to carefully move away from modelling cultural change with biological models and develop cultural models — also to make the generative inference approach an useful endevour.

Generative inference frameworks allow a straightforward application of such theoretical models to empirical data to inform on processes of social learning that are consistent with the available data.

However, how much information about those learning processes is indeed contained in specific population-level frequency data can vary from application to application.

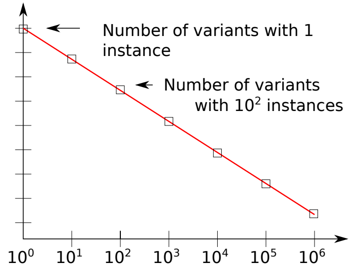
"Quality of data" problem

Data about the cultural composition of (a sample of) the population at a **large number of consecutive time points**

→ Baby name statistics

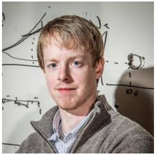
Neutral progeny distribution

Statistic: Progeny distribution recording the abundances of cultural variants which produce k new individuals (progeny) over a fixed period of time Bentley et al. (2004)



Neutral progeny distribution

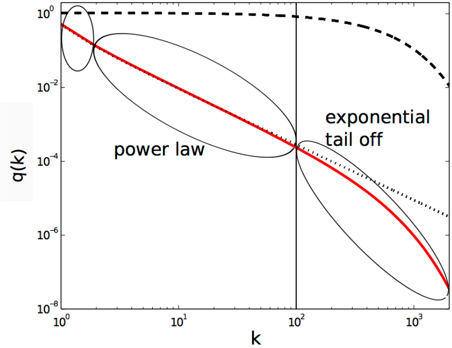
O'Dwyer and Kandler (2017)



Neutral progeny distribution

Expectation (based on a neutral birth-death process)

$$q(k) = (-1)^{k-1}\binom{\frac{1}{2}}{k}\frac{2d/b}{1+d/b}\left(\frac{4}{\frac{1}{d/b}+2+d/b}\right)^{k-1}$$



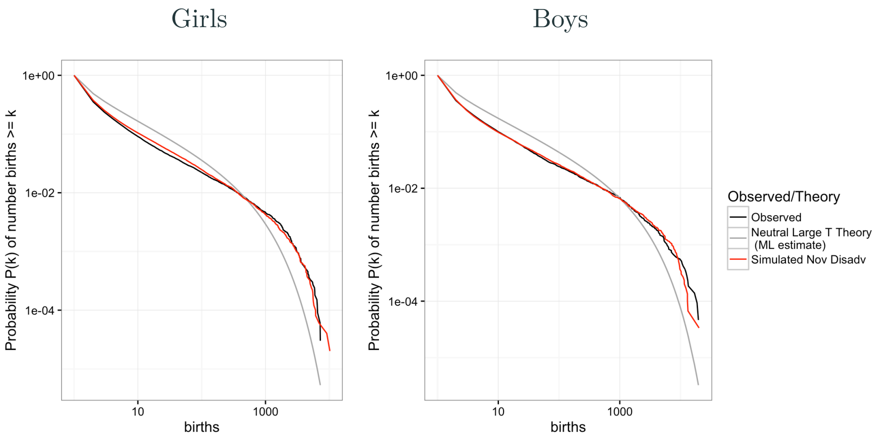
Neutral progeny distribution

$$q(k) = (-1)^{k-1} \binom{\frac{1}{2}}{k} \frac{2d/b}{1 + d/b} \left(\frac{4}{\frac{1}{d/b} + 2 + d/b} \right)^{k-1}$$

Statistical comparison: maximum likelihood estimation for d/b

$$\frac{d}{b} = \frac{K_{\text{total}}}{K_{\text{total}} - S}$$

Application to Southern Australia baby names



https://github.com/odwyer-lab/neutral_progeny_distribution

Incomplete data

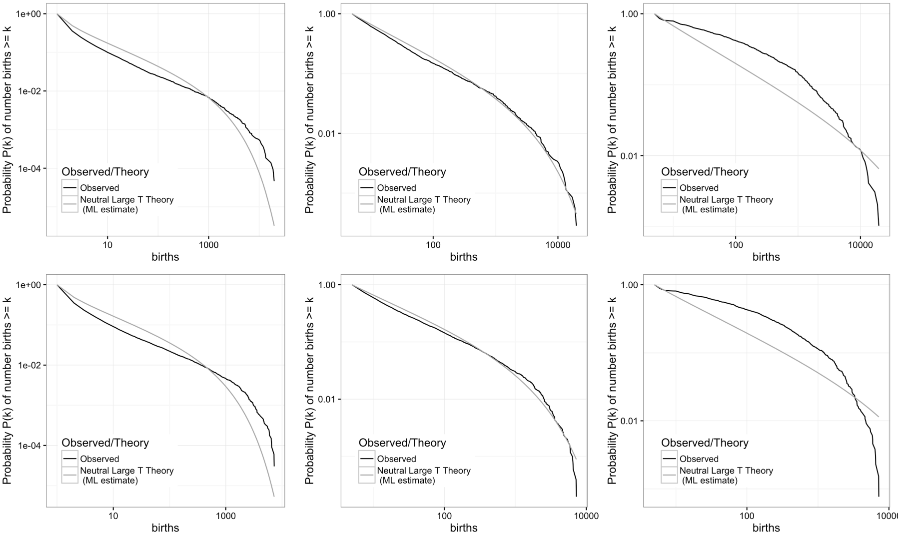
→ Many data set for registered baby names in other regions are incomplete; providing only the most popular names due to privacy considerations.

Two common ways of preprocessing cultural frequency data

- Remove all names containing < 5 instances in considered time interval (total threshold)
- Remove all names containing < 5 instances in each year (year-by-year threshold)

Incomplete data

Complete data | Total threshold | Year-by-year threshold



Incomplete data

Even if we have a lot of data omitting key parts of the data set (such as all low-frequency variants) can lead to misleading inference results.

$\rightarrow$ Important role of rare variants for inferential purposes

Conclusions

Demographic and cultural properties of the population can have a great influence on the dynamic of cultural change.

→ We need to carefully move away from modelling cultural change with biological models and develop cultural models — also to make the generative inference approach an useful endevour.

Generative inference frameworks allow a straightforward application of such theoretical models to empirical data to inform on processes of social learning that are consistent with the available data.

However, how much information about those learning processes is indeed contained in specific population-level frequency data can vary from application to application.

Thanks to you

... for listening!

Please get in touch if you have any questions!

anne_kandler@eva.mpg.de

TI(CE) LAB
Theory in Cultural Evolution