

# Reinventing Test and Trace

A Bayesian Approach For Estimating SARS-CoV-2 Setting-Specific Transmission Rates

Tim Hargreaves and Patricio Hernandez





University of Warwick  
British Conference of Undergraduate Research 2021

April 9, 2021

# Overview

1. Introduction
2. Bayesian Inference - Why and How?
3. Modelling
4. Diagnostics
5. Deployment and Scalability
6. Further Research

# Introduction - About Us

- Tim is a final-year Mathematics and Statistics student at the University of Warwick.
  -  Email: [hello@timhargreaves.com](mailto:hello@timhargreaves.com)
  -  GitHub: [THargreaves](#)
- Patricio is a second-year Economics student, also at Warwick.
  -  Email: [hdzsen.patricio@gmail.com](mailto:hdzsen.patricio@gmail.com)
  -  GitHub: [patohdzs](#)


# Introduction - Motivation

- Ineffectiveness of Test and Trace programme
  - Escalating costs of over £22 million. (National Audit Office, 2020)
  - Failure to meet testing and contact tracing targets. (iSAGE, 2020)
  - Lack of data and analysis targeted at understanding setting-specific transmission.
- Ineffectiveness of the Tier system in the UK
  - Lockdown-induced economic losses of £355 billion. (OBR, 2021)
  - Unreasonable groupings of activities with disparate transmission risks.
- **Can we design a data-driven approach for studying setting-specific transmission?**

# Introduction - Existing Work

- SIR-type models
  - Characterize epidemic evolution through dynamic compartment flows.
  - Calvetti et al. (2020), Kubinec et al. (2020), Flaxman et al. (2020).
  - Dynamic complexity increases estimate variance and error.
- Network-based modelling
  - Analyse mobility along a graph, with an overlaid SIR model.
  - Chang et al. (2020), Chang, Hradings, et al. (2020).
  - Difficult to calibrate with real data.
- Meta-analysis studies of setting-specific transmission
  - Pooled setting-specific SAR estimates from multiple scientific studies.
  - Thompson et al. (2020), Chian Koh et al.(2020).
  - Heterogeneity and practicality concerns.

# Introduction - Our Approach

- Through the T&T programme, individuals fill out a survey detailing whether they visited each one of  $P$  locations in the last week:  $\mathbf{x}_i \in \{0, 1\}^P$
- Survey data is combined with antigen test results for each individual:  $y_i \in \{0, 1\}$
- We then employ a Bayesian inference model to estimate the individual transmission rates in each of the locations included in the survey:  $\boldsymbol{\theta} \in [0, 1]^P$
- Full model description and code:  **THargreaves/reinventing-test-and-trace**

# Bayesian Inference - Why and How?

- Properties of Bayesian Inference
  - Quantifies uncertainty in parameters by treating them as random variables.
  - Generative modelling approach can perform better with fewer training samples. (Ng, 2001)
  - Incorporates experimental estimates and expert opinion through 'priors'.
- Posterior sampling done using HMC algorithm (NUTS variant).
- Two implementations:
  - Stan - Probabilistic programming language compiled in C++  $\implies$  fast!
  - TensorFlow Probability - offers support for GPU's and distributed computing setups.

# Modelling - The Base Model

- Transmission for an individual depends on:
    - Places visited.
    - Specific transmission rates of said places.
  - Transmission could occur at some other setting not specified in the survey
    - We incorporate this possibility through the base transmission risk.
  - For an individual to become infected, all we need is for transmission to have occurred in at least one setting or base transmission.
- Transmission:  $\mathbf{t}_i \sim \text{Ber}(\mathbf{x}_i \circ \boldsymbol{\theta})$
  - Base transmission:  $b_i \sim \text{Ber}(\rho)$
  - Infection:  $w_i = \mathbb{1}\{\mathbf{1}^\top \mathbf{t}_i + b_i > 0\}$



# Modelling - Resampling

- We want to relax the random sampling assumption
  - Infected individuals are more likely to get tested.
  - Symptoms, time/effort involved in testing, etc...
- Solution: Gather data from T&T as well as a random survey.
- With this framework, we can now introduce conditional testing rate parameters.
- We then 'weight down' transmission likelihood by the testing rates for infected and non-infected individuals.

- Define  $s_i = \begin{cases} 1, & \text{T\&T sample} \\ 0, & \text{otherwise} \end{cases}$
- Define also testing rates:
  - $\mathbb{P}(s_i = 1 | w_i = 1) = \gamma_+$
  - $\mathbb{P}(s_i = 1 | w_i = 0) = \gamma_-$

# Modelling - Inaccurate testing

- We also want to model test inaccuracy
    - False Positive: Test wrongly indicates infected.
    - False Negative: Test wrongly indicates not infected.
  - Solution: Introduce parameters for the true positive and true negative rates.
  - Use these to 'weight down' transmission likelihood.
  - We can incorporate 'strong priors' for TP/TN rates
    - Calibrate targeting results from UK SARS-CoV-2 test development and validation cell. (PHE Porton Down et al., 2020)
- True positive rate:  
 $\mathbb{P}(y_i = 1 | w_i = 1) = \lambda_+$
  - True negative rate:  
 $\mathbb{P}(y_i = 0 | w_i = 0) = \lambda_-$
  - Priors:  $\lambda_+, \lambda_- \sim \text{Beta}(\alpha_0, \beta_0)$
  - Calibrate  $\alpha_0, \beta_0$  to match experimental  $\bar{\lambda}$  and  $s_{\lambda}^2$ .

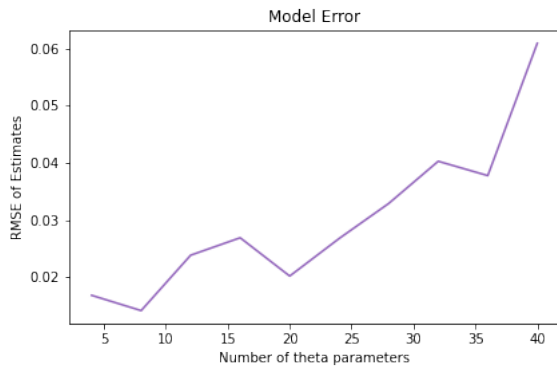
# Modelling - Hierarchical Extension

- Question: How much riskier is the Dirty Duck relative to your average pub?
    - We want to model transmission risk relative to similar settings.
  - Split similar settings into "classes".
    - E.g// fast food locales - McDonald's, Subway, Domino's, etc.
  - Model setting-specific transmission rates as deviations from the average setting class transmission rate.
- Introduce  $K$  classes.
  - Denote the class of setting  $p \in [P]$  by  $c(p) : [P] \rightarrow [K]$ .
  - Class transmission parameters:  
 $\mu \in [0, 1]^K, \sigma^2 \in \mathbb{R}_{>0}^K$
  - $\theta_p \sim \text{Logitnormal}(\text{logit}(\mu_{c(p)}), \sigma_{c(p)}^2)$

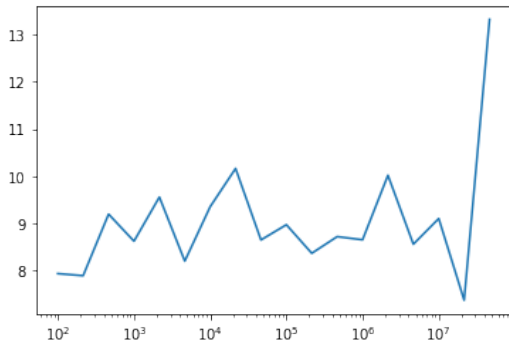
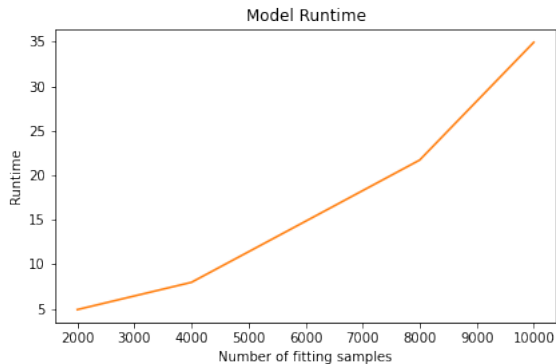
# Modelling - Interventions

- Question: Does wearing a mask have the same impact everywhere?
    - We want to model setting-specific impact of mask-wearing.
  - Ask individuals if they wear masks.
    - Capture impact on likelihood through a parameter applying only to mask-wearing individuals
  - We model the multiplier effect on the transmission odds ratio for specific classes
  - *"You are  $X$  times more likely to become infected without a mask in Setting A"*
- Define  $m_i = \begin{cases} 1, & \text{mask-wearing} \\ 0, & \text{otherwise} \end{cases}$
  - Define  $\iota \in \mathbb{R}^K$  as the class-specific mask-wearing impact.
  - $\theta_p \sim \text{Logitnormal}(\text{logit}(\mu_{c(p)}), \sigma_{c(p)}^2) + \ln(\iota_{c(p)})$

# Diagnostics - Estimation Error



# Deployment and Scalability - Runtime



# Further Research

- Modelling additional policy interventions
  - Social distancing measures, business closures, full lockdown.
- Reducing computational cost
  - Can we make more efficient code for both implementations?
- Incorporation into a hierarchical SIR
  - Use of particle filters given noise both in partial observations and dynamical system.

# References I

- Calvetti, D. et al. (2020). Metapopulation Network Models for Understanding, Predicting, and Managing the Coronavirus Disease COVID-19. *Frontiers in Physics*. <https://doi.org/10.3389/fphy.2020.00261>.
- Chang, S., Pierson, E., Koh, P.W. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 589, 82–87 (2021). <https://doi.org/10.1038/s41586-020-2923-3>
- Chang, S.L., Harding, N., Zachreson, C. et al. Modelling transmission and control of the COVID-19 pandemic in Australia. *Nat Commun* 11, 5710 (2020). <https://doi.org/10.1038/s41467-020-19393-6>
- Dillon, J.V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D.A., Patton, B., Alemi, A.A., Hoffman, M., Saurous, R. (2017). TensorFlow Distributions. *ArXiv*, abs/1711.10604.



# References II

- Flaxman, S., Mishra, S., Gandy, A. et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 584, 257–261. <https://doi.org/10.1038/s41586-020-2405-7>
- Independent SAGE. Towards an integrated find, test, trace, isolate, support (FTTIS) response to the pandemic. Jun 2020. <https://www.independentsage.org/wp-content/uploads/2020/06/IndependentSAGE-report-4.pdf>.
- Koh WC, Naing L, Chaw L, Rosledzana MA, Alikhan MF, Jamaludin SA, et al. (2020) What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate and associated risk factors. *PLoS ONE* 15(10): e0240205. <https://doi.org/10.1371/journal.pone.0240205>.
- Kubinec, R., Carvalho, L., Barceló, J., Cheng, C., Messerschmidt, L., Duba, D., Cottrell, M. S. (2020, April 1). Fear, Partisanship and the Spread of COVID-19 in the United States. <https://doi.org/10.31235/osf.io/jp4wk>.

# References III

- National Audit Office. (2020). The government's approach to test and trace in England – interim report. <https://www.nao.org.uk/press-release/the-governments-approach-to-test-and-trace-in-england-interim-report/>.
- Ng, A., Jordan, M.(2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. Advances in Neural Information Processing Systems.
- Office for Budget Responsibility. (2021). Overview of the March 2021 Economic and fiscal outlook. <https://obr.uk/overview-of-the-march-2021-economic-and-fiscal-outlook/>.
- Riddell, A., Hartikainen, A., Carter, M. (2021). PyStan (3.0.0). <https://pypi.org/project/pystan>.
- Thompson, H. et al. (2021). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Setting-specific Transmission Rates: A Systematic Review and Meta-analysis. Clinical Infectious Diseases. <https://doi.org/10.1093/cid/ciab100>.