Project Plan
Seth Temple
June 9, 2021

- Methods
    - I present and prove Lemma 1 and Lemma 2 in Stockdale et al. (2019). For Lemma 1, I take advantage of exponential racing as presented in Durrett (1999). This probabilistic argument illustrates how to extend to Gamma infectious periods (Lemma 4).
    - I exclude a complete proof for Lemma 4 in Stockdale et al. (2019) and Stockdale (2019). Extending pairwise general moments to Gamma infectious periods involves the same proof strategies as those of Lemma 1 for exponential infectious periods. I provide a high-level argument for this extension in my manuscript. Moreover, I have worked these out on paper, resulting in me correcting an index subscript in their paper.
    - To prove Lemma 3, I appeal to Theorem 2.1 of Barbour and Eagleson (1985). This application is shown in Appendix A. I do not prove Theorem 2.1 of Barbour and Eagleson (1985) but comment on its additional dissociation (independence) assumption. I have also corrected a constant $(2n - 1)$ to $(4n - 5)$, which does not influence the results. Finally, I show in a simulation study that this normal approximation provides no runtime improvement relative to an exact distributional result.
    - I frame the stochastic epidemic model as a continuous-time Markov chain, specifically as a Markov jump process. This framing reconciles the proof strategies with an event-driven epidemic simulator that I devised. I had to devise this algorithm to simulate epidemics because Stockdale et al. (2019) say nothing about simulating epidemics.

- Simulation Studies
    - I developed an R package sdtemple/pblas to conduct simulation studies and real data analyses. Whereas Stockdale et al. (2019) share a few scripts (jessicastockdale/PBLA) as examples of their real data analyses, my package comprehensively presents six pair-based likelihood approximations, is entirely in the R scripting language, and requires no dependencies. It is also fast.
    - Using my R package, I reproduced the simulation studies in the supplement of Stockdale et al. (2019).
    - I conducted **additional simulation studies** to (i) measure the computational runtime of these methods, (ii) test when the methods fail, (iii) investigate the consistency of estimators, and (iv) assess the effects of case underreporting. I indicate in (iii) that MLEs from PBLA are **not consistent estimators** of the true parameters.

- Data Analyses
  - I exactly reproduced the PBLA MLEs for an Ebola virus epidemic in West Africa.
  - I built an MCMC sampler with random walk proposals to replicate an analysis of common cold cases in a remote Atlantic island (Tristan da Cunha).
  - I could not access the 2001 data on foot-and-mouth disease in the United Kingdom. I replaced this dataset with a rabies viral epidemic in Bangui, Central African Republic. I investigated a spatial component to this epidemic, tested the sensitivity of the analysis to *N*, the total population size, and addressed case underreporting with two MCAR adjustments.

- Courses
  - STAT 581-3: Advanced Statistical Inference
  - STAT 570-1: Advanced Regression Methods
  - STAT 559: Measure Theory
  - STAT 516: Stochastic Modeling
  - STAT 512-3: Statistical Inference
  - BIOST 533: Theory of Linear Models
  - BIOST 550-551: Statistical Genetics
  - BIOST 581: Statistical Genetics Seminar
  - GENOM 540: Computational Molecular Biology