Fixed-time descriptive statistics underestimate extremes of epidemic curve ensembles
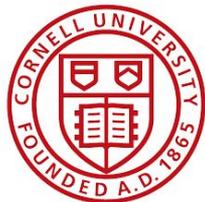
— and how to fix it?

Jonas L. Juul – Cornell University
✉ jjuul@cornell.edu   🐦 @jonassjuul
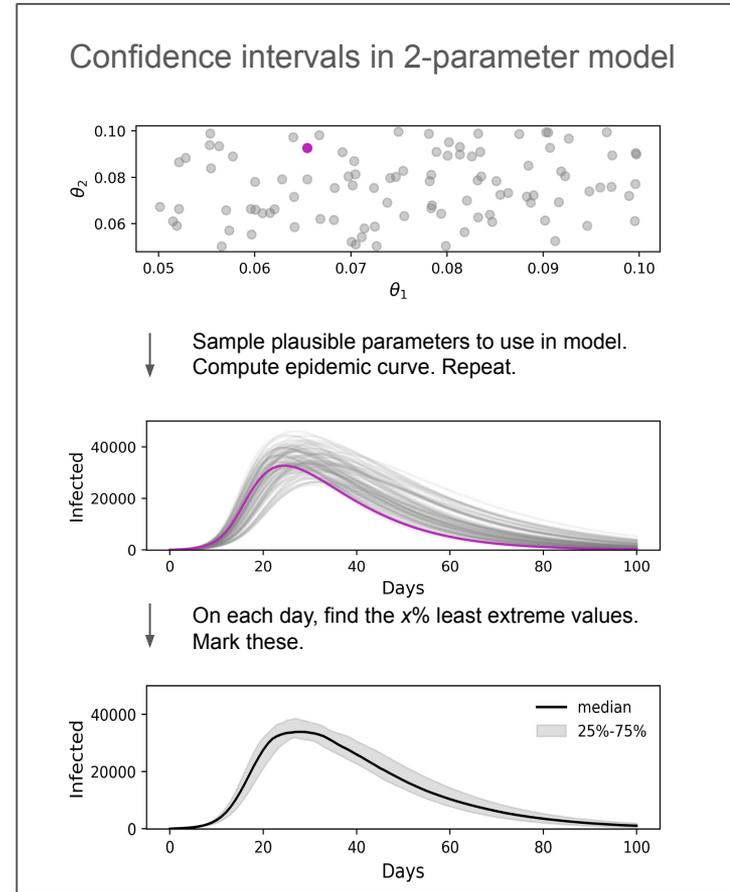
Center for
Applied Mathematics

The topic:
Uncertainty in epidemiological projections

Why?

- Models take input parameters that are known to some (limited) precision.
- Locking down/reopening society are calculated risks. Uncertainty on projections must be well-communicated.

How is uncertainty usually communicated?

- *See illustration.*

Confidence intervals in 2-parameter model

Sample plausible parameters to use in model. Compute epidemic curve. Repeat.

On each day, find the *x*% least extreme values. Mark these.

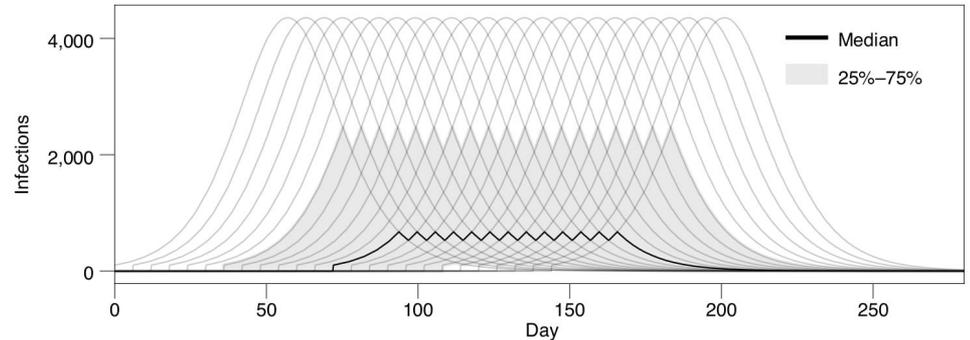Juul et al., *Nature Physics* **17,** 5–8 (2021)

The Tragic Tale of Transmithaca

What is the problem?

- For COVID-19, many countries locked down to prevent hospitals from overflowing.
- Current summary of uncertainty is not fit for communicating risks of such extremes.

Transmithaca

1. People on the island of Transmithaca know a virus will arrive. They know every parameter with perfect precision, but not *when* it arrives.
2. What is the expected peak number of infections per day?
3. Projection suggests < *3000 infections per day* although every curve exceeds *4000* infections.

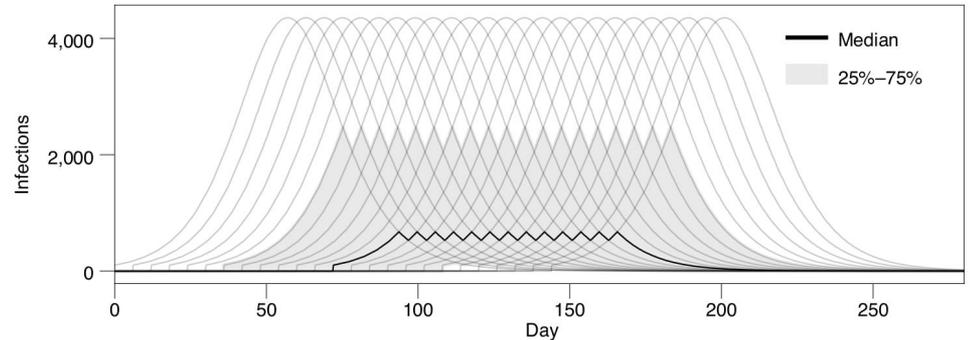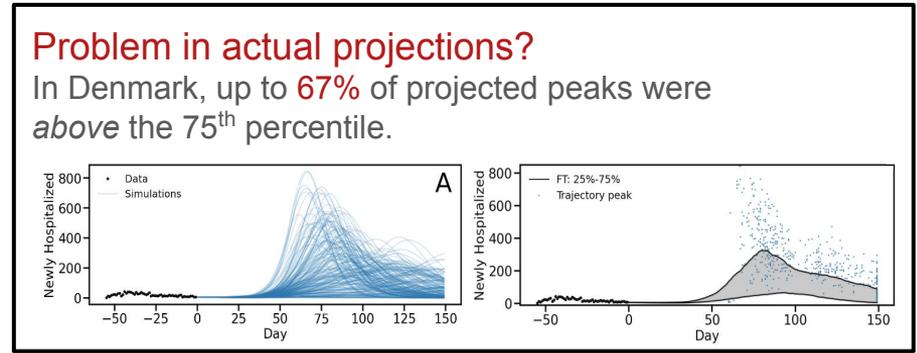Juul et al., *Nature Physics* **17,** 5–8 (2021)

The problem:
The Tragic Tale of Transmithaca

What is the problem?

- For COVID-19, many countries locked down to prevent hospitals from overflowing.
- Current summary of uncertainty is not fit for communicating risks of such extremes.

Transmithaca

1. People on the island of Transmithaca know a virus will arrive. They know every parameter with perfect precision, but not *when* it arrives.
2. What is the expected peak number of infections per day?
3. Projection suggests < *3000 infections per day* although every curve exceeds *4000* infections.

Problem in actual projections?
In Denmark, up to 67% of projected peaks were *above* the 75th percentile.

Juul et al., *Nature Physics* **17,** 5–8 (2021)

The solution:
Statistics for whole curves

Solution 1: Curve-based descriptive statistics

- Rank centrality of whole curves instead of single points in time. Draw envelope of most central. *[There are several ways to rank curves by centrality]*

Solution 2: Estimate likelihood of scenarios

1. Directly estimate likelihood of scenarios of interest from ensemble.
   a. Example: Probability of at least 200 newly hospitalized for at least 2 consecutive days is above 0.50 because this scenario took place in more than 50% of the simulations.

✉ jjuul@cornell.edu   🐦 @jonassjuul