Chris Lab Notebook

Chris Hoover

5/11/2020

TODOs

5/11/20 Clarify source of iterations/updating with Josh
5/11/20 Get R_0 equation as function of model parameters from NextGen matrix

5/11/2020

Goals:
* Familiarize with model
* Code model in STAN and try some initial fits to get idea of runtimes, outputs
* Start brainstorming potential improvements to model

Package/Model functionality

R/InputsFromSpreadsheets.R

Contains functions to read user inputs and translate them into model inputs, generate parameter sets to sample from. Also main wrapper function that users run to generate report: `CredibilityIntervalFromExcel`

R/CredibilityInterval.R

contains functions to run model, check which model runs "fit" input data (e.g. which parameter sets generate predictions that agree with input data)

R/CombinedModels.R

contains functions to run seir model

R/CreateOutputs.R

Takes model posteriors and generate plots of projections, posterior distributions and comparison to prior distributions

Overall process

1. User calls `CredibilityIntervalFromExcel` which processes data in excel spreadsheet and then calls `CredibilityInterval`

2. `CredibilityInterval` does some more preprocessing of model inputs and then calls `RunSim1`

3. `RunSim1` calls `RunSim` which uses `FitSEIR` to generate reasonable estimates of when the outbreak started based on comparison of model generated outputs to observed data done with `CalcError` function. Once FitSEIR has found a best start data for each parameter set, `Seir` is run.

4. The output of `RunSim1` comes from running `Seir` which produces model estiamtes from the input parameter estimate and best guess start date. `InBounds` is then called to determine if the output of `Seir` is within the user provided uncertainty bounds of the observed data.
   ** There seems to be some sort of iteration updating going on that I can't quite figure out

Test Run

```
file.copy(system.file("extdata", "SF-April13.xlsx", package = "LEMMA", mustWork = TRUE), "example.xlsx")

LEMMA::CredibilityIntervalFromExcel("example.xlsx")
```

Works for the most part except for a fairly obscure ggplot error

Implementing in STAN

Priors

Need to explicitly define distributions of priors in STAN, which can mostly be done from the same framework already built into LEMMA inputs.

```
# Some Rstan setup options
rstan_options(auto_write = TRUE)
options(mc.cores = parallel::detectCores())

# Rstan model
cat(
  '
  functions{}
  '
)
```

```
##
##   functions{}
##
```

Potential improvements (running list)

Continuous time, stochastic model versions

Erlang distributed latent period, hospitalization length

Make some priors endogenous (e.g. non-adjustable), focus users on things most likely to vary by region/area/population

- Population Size

- Start date of projection

- End date of projection

- Percent infected who are hospitalized

- Average hospital length of stay

- Timing and effectiveness ($\mathcal{R}_1$) of interventions