THESIS PROPOSAL

HMM modelling for the spread of the SARS-CoV-2

Martin Beneš

Supervisor: Krzysztof Bartoszek

5 Feb 2021

Presentation structure

- Data description
- Research questions
- Methods to solve the questions
- Possible bottlenecks
- Similar research

Covid-19 statistics

- Daily, weekly
- Attributes
  - **Tests** = number of performed tests (RT-qPCR + antigen)
  - **Confirmed cases** = number of newly positively tested people
  - **Deaths** = number of deceased on/with Covid-19
  - **Hospitalized** = number of new hospitalized
  - **Recovered** = number of recovered/released from hospital
- Country-wise, regions, municipalities

Covid-19 statistics: Czechia

- https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19
- **Country:** RT-qPCR + antigen tests
- **District:** deaths, tests, hospital capacities, hospital stock states
  - Per age group: incidence, prevalence, hospitalized, vaccinated
  - Cases with age and gender: confirmed, deaths
- **Municipality:** confirmed

Covid-19 statistics: Poland

- https://www.gov.pl/web/koronawirus/pliki-archiwalne-powiaty
- https://twitter.com/MZ_GOV_PL
- **Country:** tests, recovered, hospitalized, quarantined
- **Region/municipality:** confirmed, deaths,

Covid-19 statistics: Sweden

- https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/
- https://scb.se/om-scb/nyheter-och-pressmeddelanden/overdodligheten-fortsatter-att-sjunka-efter-toppen-i-april/
- **Country:** deaths, icu, confirmed (daily)
- **Region:** icu (weekly), vaccines, tests - antibody
- **Municipality:** confirmed, deaths (weekly)

Covid-19 statistics: Italy and others

- https://covid19datahub.io/
- Unified datahub for Covid-19 data
- Italy + Switzerland:
  - **Municipality:** deaths, confirmed, tests, recovered, hospitalized, icu
- Implemented: https://github.com/covid19datahub/COVID19/blob/master/inst/extdata/src.csv

Demographical statistics

- **Mortality**
  - https://ec.europa.eu/eurostat/web/products-datasets/-/demo_r_mweek3
- **Population**
  - https://ec.europa.eu/eurostat/web/products-datasets/-/tgs00096
- **Area, …**

Research questions

- **Data:** To what extent are the collected data reliable?
- **Model:** What are the distributions of parameters of Covid-19?
  - Incubation period
  - Infection fatality ratio
  - Reproduction number
  - Duration of disease
- **Simulation:** Are the reported statistics projected in the results?
- **Simulation:** Are there visible patterns or similarities between regions?
- **Simulation:** Are the introduced restrictions in the region/country projected in the numbers yielded by the simulation?

Data: To what extent are the collected data reliable?

- Evaluate the sampling method.
- Is number of performed tests sufficient?
- Analyze the data for discrepancies.

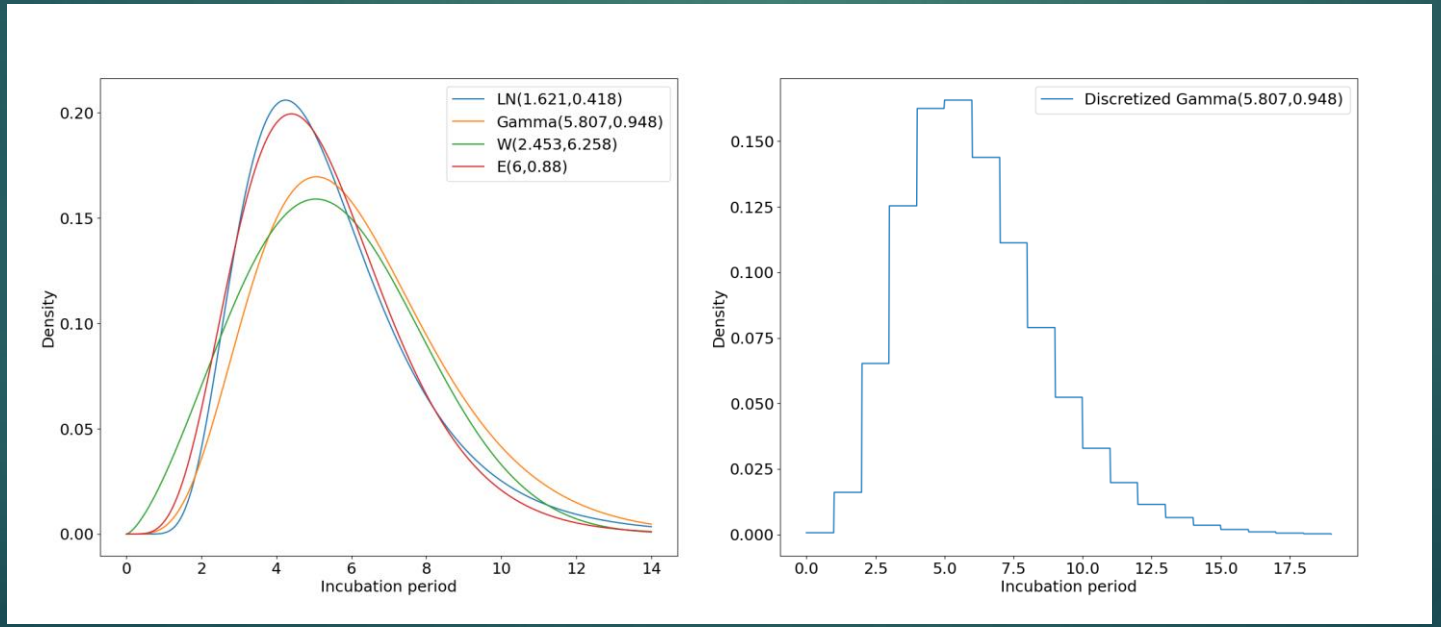
Figure 21. P-values for equal ratio t-test (eq. 10).

Day	Country		
	Czechia	*Poland*	*Sweden*
Monday	0.581	0.001	0.429
Tuesday	0.496	0.06	0.088
Wednesday	0.784	0.112	0.731
Thursday	0.375	0.181	0.924
Friday	0.298	0.764	0.507
Saturday	0.112	0.737	0.394
Sunday	0.294	0.044	0.947

Source: own unpublished research project

Model: What are the distributions of parameters of Covid-19?

- SIR* model
  - Parameters a,b,c,d needed
- Objective of other papers

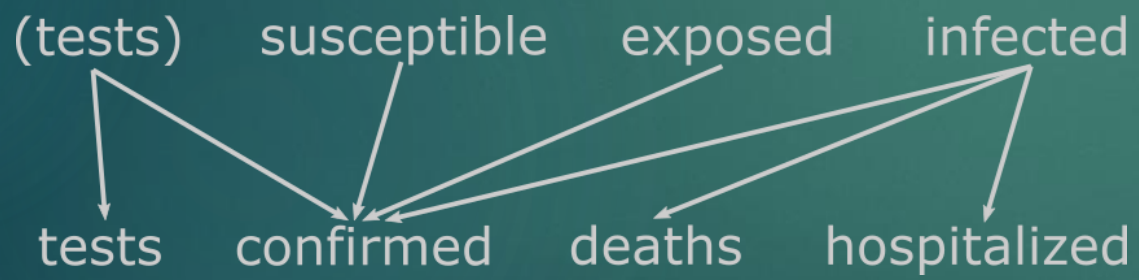


Source: [10.7326/M20-0504](10.7326/M20-0504)

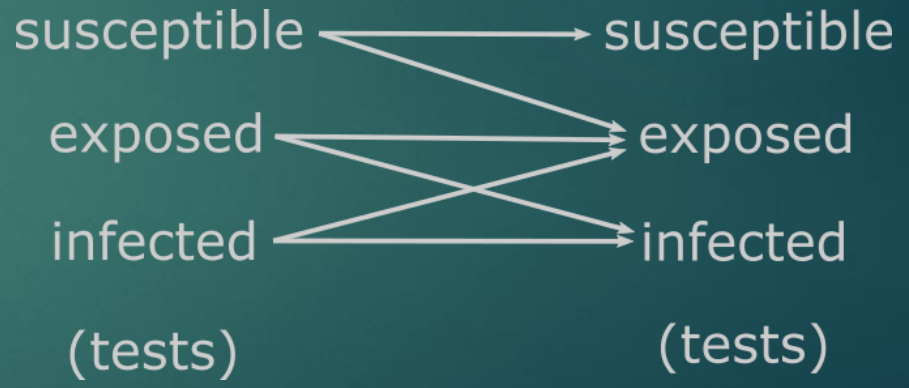
Model: What are the distributions of parameters of Covid-19?

- HMM model
  - Transition/emission matrices based on SIR*

Emission matrix structure



Transition matrix structure

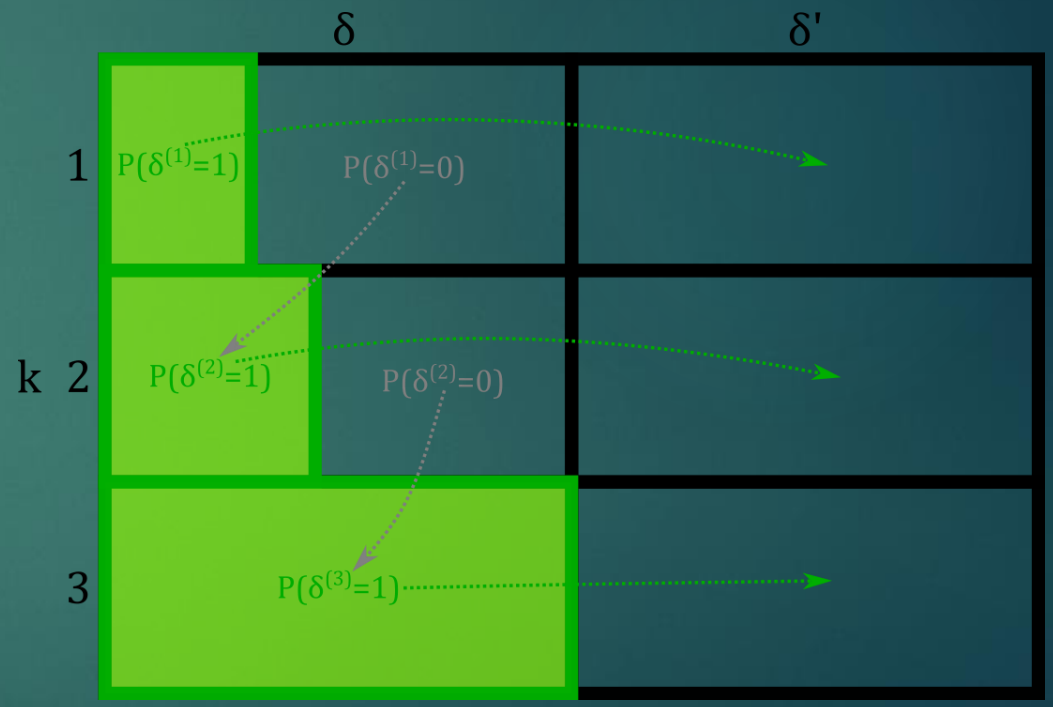


Model: What are the distributions of parameters of Covid-19?

- HMM model

$$P(\delta^{(1)} = 1) = P(\delta = 1)$$
$$P(\delta^{(k)} = 1 | \delta^{(k-1)} = 0, \ldots, \delta^{(1)} = 0) = \frac{P(\delta = k)}{\prod_{i=1}^{k-1} P(\delta^{(i)} = 0)} \quad (2.4)$$

Source: Master thesis



Source: Master thesis

Simulation: Are the reported statistics projected in the results?

- Do the reported statistics describe similar situation as the simulation results?
- Are confirmed cases a good measure of the situation?
- Is there a visible connection with testing strategy?

Simulation: Are there visible patterns or similarities between regions?

- Comparison of infected.
- Possible usage of EM to drop less certain parameters.

Simulation: Are the restrictions projected in the simulated data?

- Evaluation of the situation in different regions in the context of local implemented restrictions.
- Evaluation of restriction efficiency.
- **Data of restrictions**: media, official websites, Covid-19 Data Hub

Bottlenecks

- **Symptoms / asymptomatic**
  - Different R0 number distribution (less cautious)
  - Different probabilities of being tested
- **Non-permanent immunity**
  - Additional connection of R → S
- Infectious **incoming from abroad**
  - Additional additive term in the equation
- **Vaccination**
  - Lower susceptibles
  - Problem of data
- **Test errors**
  - Test is wrong with a certain probability (specificity + sensitivity)
- **Infectiousness / symptoms**
  - Not the same

$$\frac{dS}{dt} = -aSI$$

$$\frac{dE}{dt} = aSI - cE$$

$$\frac{dI}{dt} = cE - bI - dD$$

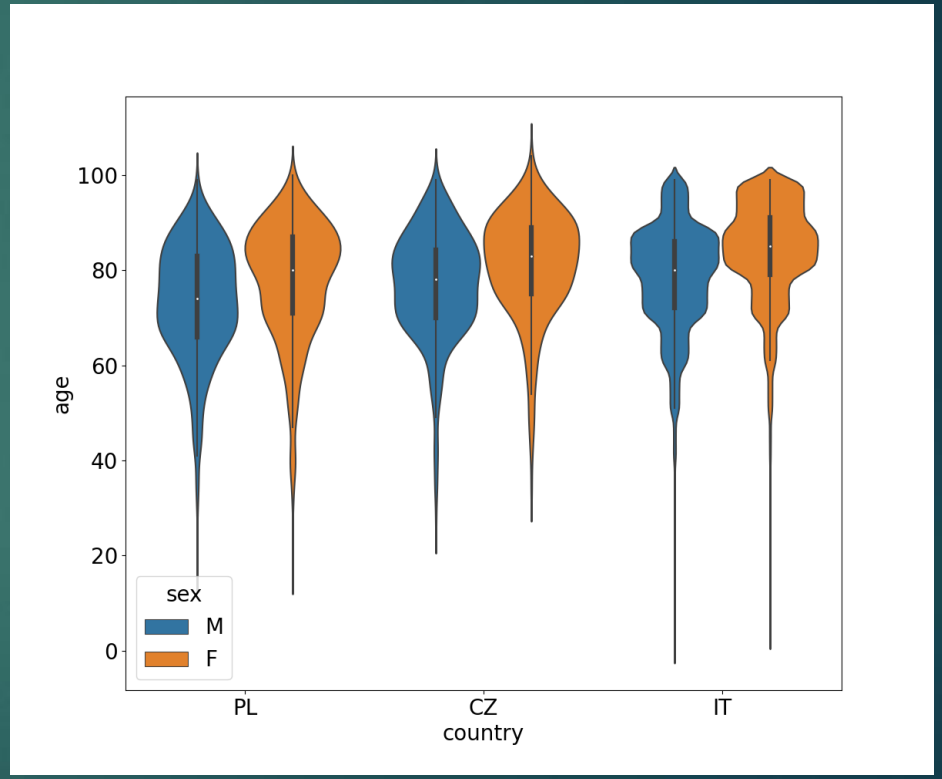
$$\frac{dR}{dt} = bI$$

$$\frac{dD}{dt} = dD$$

Source: Master thesis

Similar research

- Similar research already exists
  - From second half of the year 2020
- Parameters per population, not per-age



Source: Master thesis

Similar research

- **Tracking the early depleting transmission dynamics of COVID-19 with a time-varying SIR model**
- https://www.nature.com/articles/s41598-020-78739-8
- Compartment model
- Data from Malaysia
- Latent variable is R0

Similar research

- **COVID-19 dynamics in an Ohio prison**
- https://doi.org/10.1101/2021.01.14.21249782
- Compartment model
- Measured in Ohio prison, where was outbreak
- Might include very different parameter estimates
- *Not reviewed preprint*

Similar research

- **Symptom-based testing in a compartmental model of Covid-19**
- https://doi.org/10.1101/2020.10.11.20211037
- Compartment model
- Does not specify what data is used
- *Not reviewed preprint*

Thank you for attention.

Martin Beneš

5 Feb 2021