A new approach to modelling pandemic

Andrea Dapor
INVERENCE

Forecast Hub (ECDC), 13th April 2021

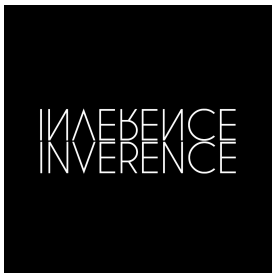


table of contents

1 general idea: our concept in modeling disease

Typical SEIR models in epidemiology require a **deterministic knowledge** of contagion and other quantities, as well as a **complete understanding of the dynamics** and the interaction between the various compartments, generally mediated via **time-independent parameters** (e.g., constant reproductive number). However, in real life:

- our observations are prone to error (e.g., number of confirmed cases vs real number of cases)
- our understanding of the system is incomplete (e.g., how does climate affect the progression of pandemic?)
- we see that the parameters of the model do in fact change in time

Ideally, we would have perfect observations and perfect models. In absence of that, what can we do?

In our approach, we cope with these issues by describing the observed contagion and deaths in terms of a **stochastic variables** which obey a probabilistic model whose parameters may also be stochastic variables:

$$x_t = n_t + f_t \qquad n_t = \frac{\sum_i \theta_{i,t} B^i}{\sum_j \varphi_{j,t} B^j} a_t \qquad B^k a_t := a_{t-k}$$

The noise component absorbs all the unknown effects, while the deterministic part contains the known inputs. Both parts may contain parameters that are themselves stochastic variables, obeying their own probabilistic equations.

2 additive model of contagion

such that $\sum_i \phi_i = 1$

$$x_t = r_t \sum_{i=1}^{p} \phi_i x_{t-i} + a_{x,t}$$

effective reproductive number

white noise

to see that r_t is really the reproductive number, set the noise to 0, and consider the simple case where the sum consists of a single term:

$$x_t = r_t x_{t-1}$$

This means that r_t is the average number of people that get infected at time t by a person who got infected at time t-1. More generally, the number of people that get infected at time t by a person who got infected at time t-i is $r_t \phi_i$.

The reproductive number obeys its own probabilistic equation:

$$r_t = n_{r,t} + f_{r,t}$$

3 multiplicative model of contagion

$$x_t = r_t \sum_{i=1}^{p} \phi_i x_{t-i} + a_{x,t} \quad \Longrightarrow \quad x_t = r_t \left(\prod_{i=1}^{p} x_{t-i}^{\phi_i} \right) \exp(a_{x,t})$$

weighted arithmetic mean

$$\bar{x}_t = \sum_i \phi_i x_{t-i} \quad \longrightarrow \quad \tilde{x}_t = \prod_i x_{t-i}^{\phi_i}$$

weighted geometric mean

Taking the logarithm, we obtain a linear equation:

$$\ln(x_t) = \ln(r_t) + \sum_{i=1}^{p} \phi_i \ln\left(x_{t-i}\right) + a_{x,t}$$

so $\ln(r_t)$ can be treated as a stochastic input. It satisfies its own equation which, once plugged in, gives

$$\sum_{i=0}^{p} \phi_i \ln(x_{t-i}) = n^*_{x,t} + f_{r,t}$$

3 multiplicative model of contagion

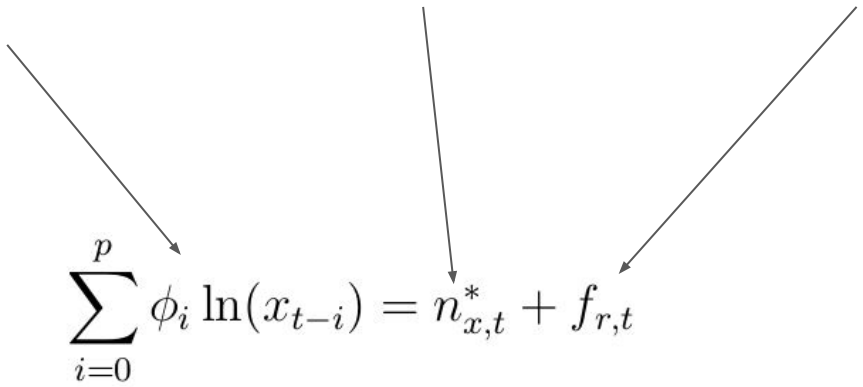
This formula hides the complexity of the model (in particular, the time-dependent reproductive number has disappeared), allowing for a simple estimation of the parameters.

Intuitive interpretation:

self-interaction: how the past values of $\ln(x_t)$ affect the current value

noise: accounting for unknown effects in $\ln(x_t)$ and $\ln(r_t)$

deterministic and stochastic inputs to $\ln(r_t)$



$$\sum_{i=0}^{p} \phi_i \ln(x_{t-i}) = n^*_{x,t} + f_{r,t}$$

4 model of deaths

such that $\sum_i \omega_i = 1$

For consistency with the contagion, we describe death also via a multiplicative model:

noise

$$y_t = g_t \left(\prod_i x_{t-i}^{\omega_i} \right) \exp(a_{y,t})$$

Deaths are proportional to a weighted geometric average of cases, with proportionality factor g_t. The interpretation of g_t is thus of **average lethality** of the virus. This is given by a stochastic part h_t and a deterministic part L_t, the weighted average over the population groups (e.g., age groups):

$$g_t = h_t L_t$$

stochastic

deterministic

$$L_t = \sum_b l_b \lambda_{b,t}$$

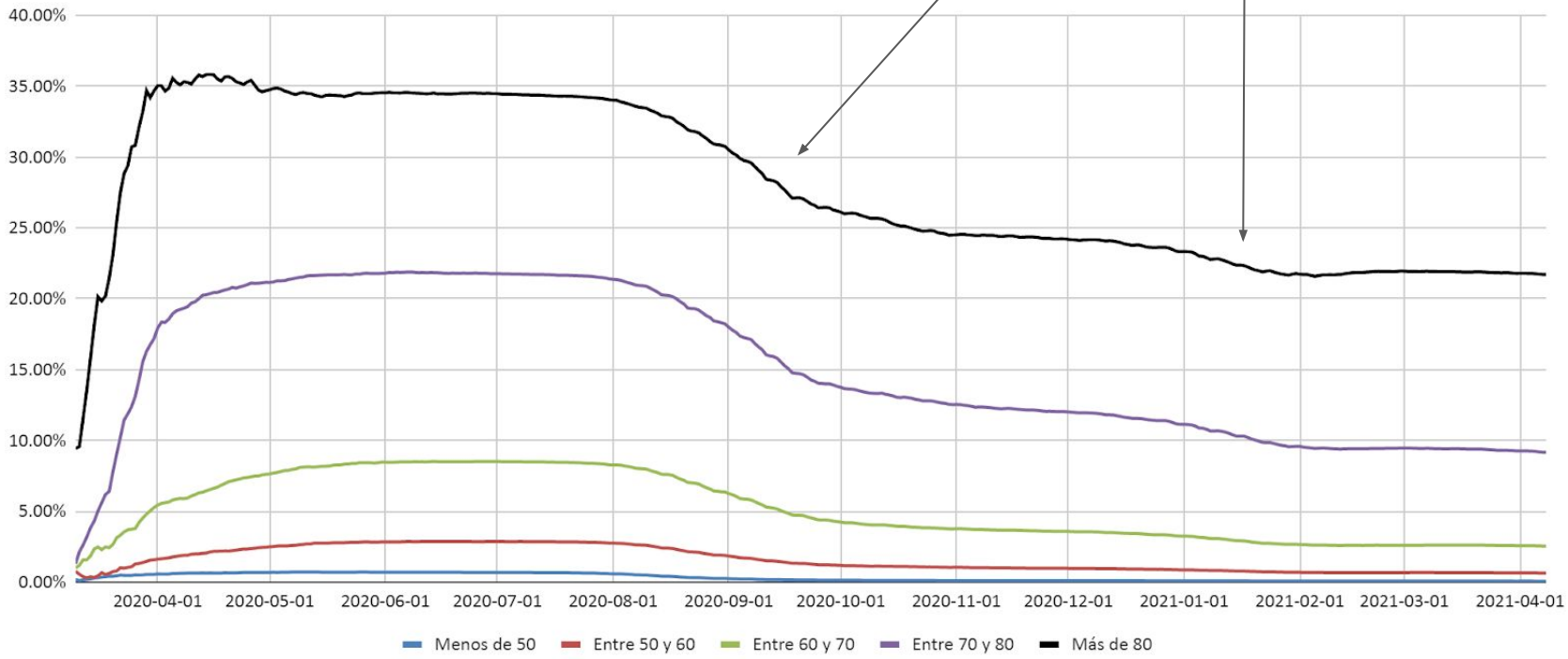
percentage of population which belongs to group b and which is susceptible

lethality for group b

4 model of deaths

What causes this dynamical behavior? We can make hypotheses and construct a deterministic model, or simply encode it in h_t.

Evolution of the lethality for different age groups (Madrid):



4 model of deaths

Absorbing h_t in the noise and taking the logarithm:

$$\ln(y_t) = n^*_{y,t} + \ln(L_t) + \sum_i \omega_i \ln(x_{t-i})$$

noise: accounting for unknown effects in $\ln(y_t)$

lethality (weighted average of groups)

input from contagion

5 vaccination strategies for Covid19

In presence of vaccination, the reproductive number acquires a deterministic component given by the "susceptibility":

stochastic

$$r_t = r_t^* \alpha_t$$

deterministic

Susceptibility α_t may or may not encodes also natural immunity (if we have a model for it), but we assume only immunity due to vaccination (natural immunity is then encoded in r_t^*). We can write

$$\alpha_t = 1 - \frac{1}{P} \sum_m \sum_b C_{m,b}(B) v_t^{(m,b)}$$

number of people in group b vaccinated at time t with vaccine m

total population

sums over all vaccines m and population groups b

transfer function of vaccine m on population group b

Population groups:

A. health workers age 18-55
B. general population of age 80+
C. general population of age 56-79
D. general population of age 18-55

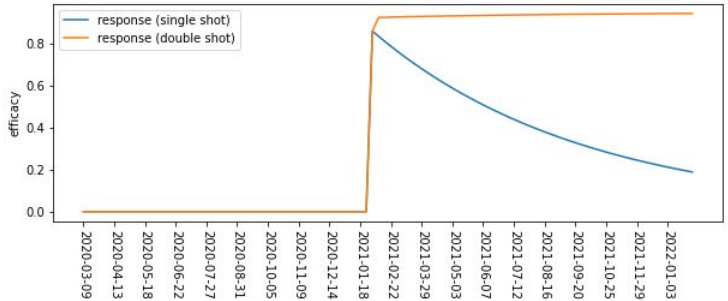
5 vaccination strategies for Covid19

The function $C_{m,b}(B)$ of operator B contains the effect of vaccine m on population group b. The simplest example is if the vaccine gives immediate and infinite immunity of $\gamma_{m,b} \in [0,1]$: then we simply have $C_{m,b}(B) = \gamma_{m,b} \sum_{i=1}^{\infty} B^i$

$$\alpha_t = 1 - \frac{1}{P} \sum_m \sum_b C_{m,b}(B) v_t^{(m,b)}$$

In our Covid19 model for the Spanish Health Department, we considered more realistic functions, with an immunity of $\delta^{(1)}$ two weeks after the first shot, which exponentially decays to $\delta^{(y)}$ after 52 weeks, and is brought up to $\delta^{(2)}$ after the second shot.

vaccine m	group b	$\delta^{(1)}$	$\delta^{(2)}$	$\delta^{(y)}$
Pfizer/ Moderna	A	0.86	0.95	0.20
	B	0.86	0.95	0.20
	C	0.86	0.95	0.20
	D	0.86	0.95	0.20
Astra Zeneca	A	0.60	0.70	0.20
	B	-	-	-
	C	0.30	0.50	0.20
	D	0.60	0.70	0.20



Population groups:

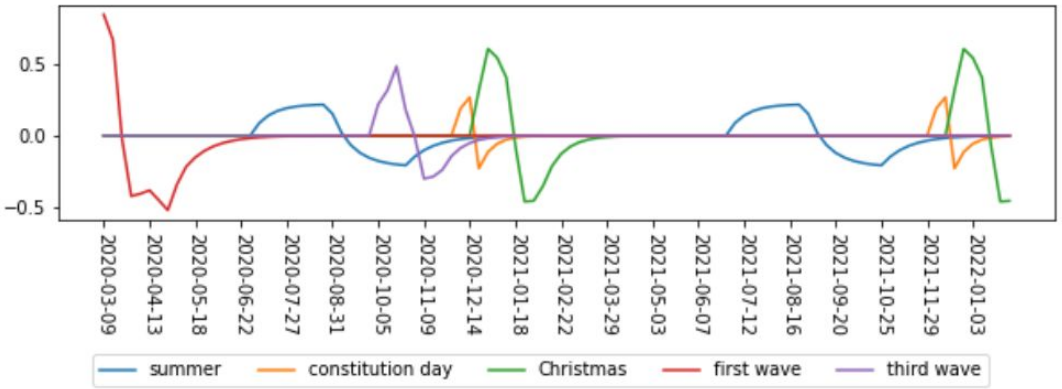
A. health workers age 18-55
B. general population of age 80+
C. general population of age 56-79
D. general population of age 18-55

5 vaccination strategies for Covid19

Summary of Covid19 model:

$$\begin{cases} (1-B)\ln(x_t) = \dfrac{1-\theta_1 B-\theta_2 B^2}{1-B}a_{x,t} + \ln(\alpha_t) + f_{r,t} \\ \ln(y_t) = \dfrac{1}{(1-B)(1-\phi B)}a_{y,t} + \ln(L_t) + (\omega_0+\omega_1 B+\omega_2 B^2)\ln(x_t) \end{cases}$$

where $f_{r,t}$ is a sum of transfer functions acting on deterministic inputs accounting for holidays (summer, Christmas and the Spanish Constitution Day) as well as the observed first and third waves.



The free parameters of the model (to be estimated via MCMC):

- θ_1, θ_2
- parameters in $f_{r,t}$
- ϕ
- ω_0 , ω_1 , ω_2
- standard deviations σ_x and σ_y

Population groups:

A. health workers age 18-55
B. general population of age 80+
C. general population of age 56-79
D. general population of age 18-55

5 vaccination strategies for Covid19

Based on vaccines availability, we considered 6 strategies for Pfizer/Moderna vaccine and 3 for AstraZeneca vaccine, for a total of 18 scenarios:

Pfizer/Moderna:

a1. single dose, group order: (A, B, C, D)

a2. single dose, group order: (B, A, C, D)

a3. double dose (3 weeks apart), group order: (A, B, C, D)

a4. double dose (3 weeks apart), group order: (B, A, C, D)

a5. double dose (6 weeks apart), group order: (A, B, C, D)

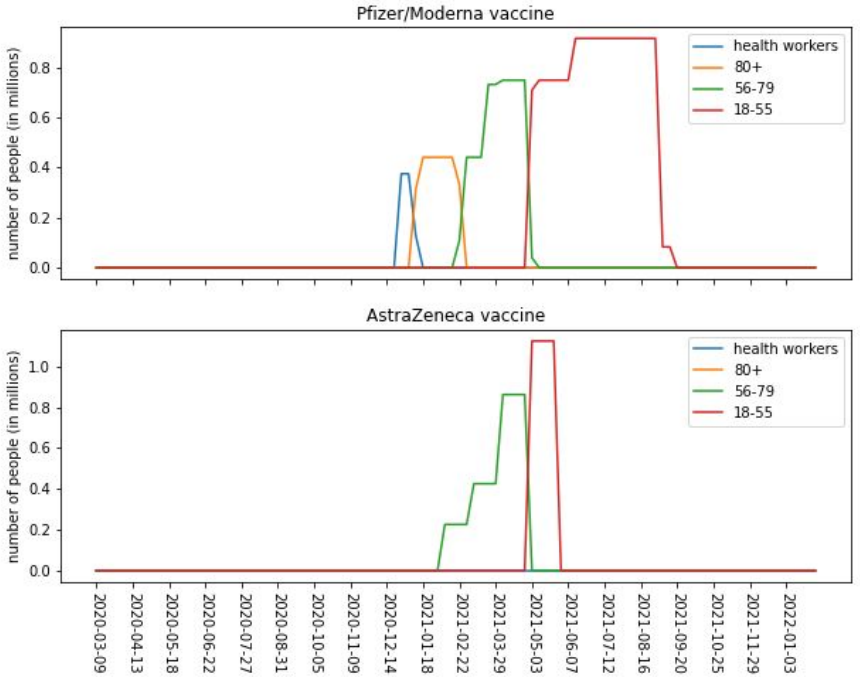
a6. double dose (6 weeks apart), group order: (B, A, C, D)

AstraZeneca:

b1. double dose (4 weeks apart), group order: (A, D)

b2. double dose (4 weeks apart), group order: (A, D, C)

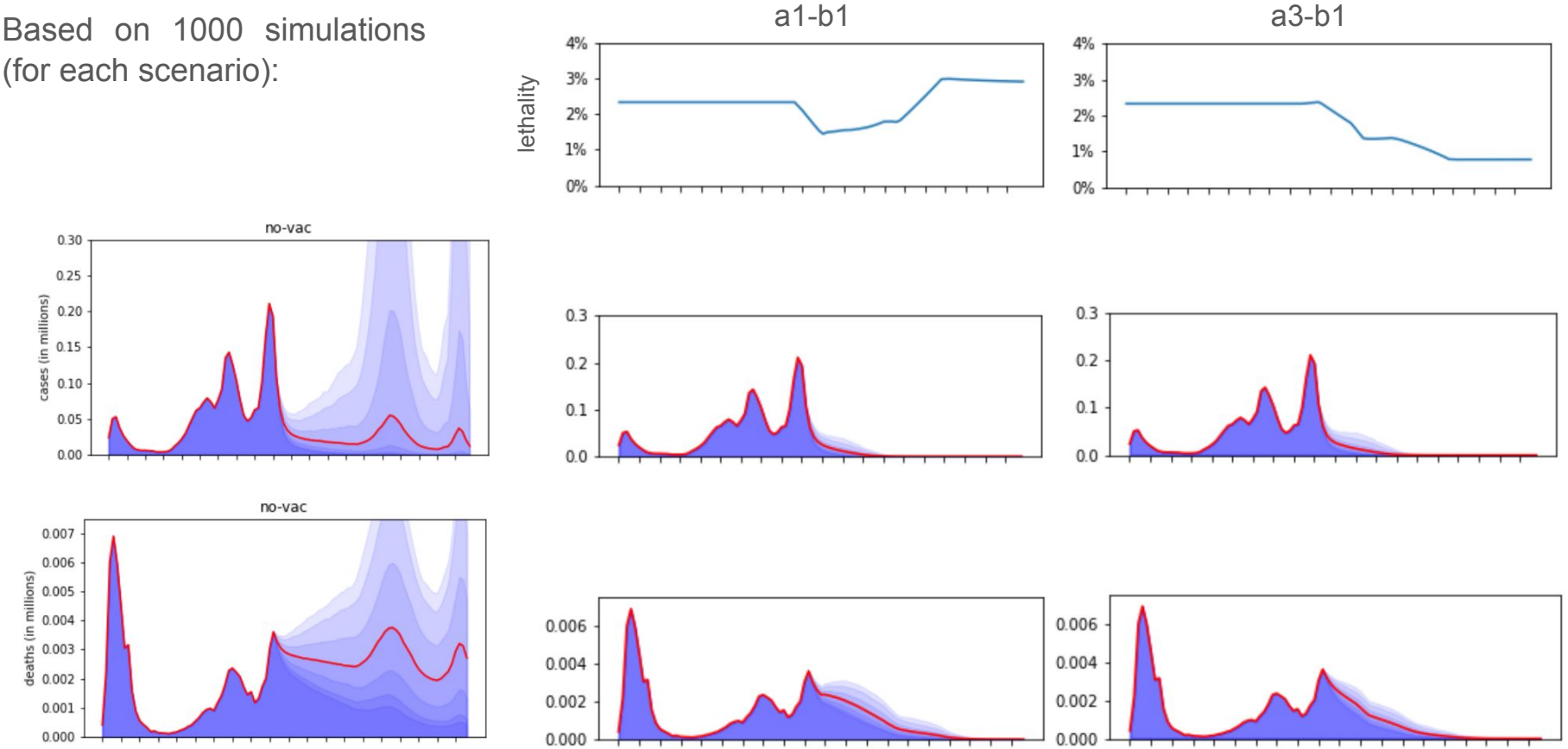
b3. double dose (4 weeks apart), group order: (C, A, D)



Example: vaccination process in scenario a1-b3

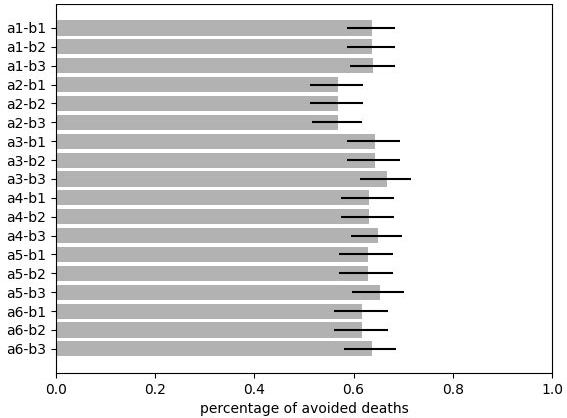
5 vaccination strategies for Covid19

Based on 1000 simulations (for each scenario):



5 vaccination strategies for Covid19

avoided deaths after 36 weeks:



	b1	b2	b3
a1	0.109	0.109	0.109
a2	0.117	0.117	0.117
a3	0.109	0.109	0.106
a4	0.110	0.110	0.108
a5	0.110	0.110	0.108
a6	0.111	0.111	0.110

expected cumulative deaths after 1 year (no-vac is 0.232):

All strategies produce similar results, but the comparison with the no-vac case shows the obvious benefits of the vaccination: not only the number of deaths is reduced even beyond 60%, but the uncertainty is also greatly reduced, since the most extreme scenarios become very unlikely.

6 conclusions

We have developed a framework for models of contagion and death based on **stochastic variables** obeying equations with **time-dependent stochastic parameters**. This allows to cope with the uncertainty coming from

- errors in historical observations, statistically censored information
- incomplete theoretical understanding

We illustrated the framework in the case of Covid19 pandemics, where we forecasted the value of observables in several vaccination scenarios.

While already good, these models are a work in progress, and can be improved as our theoretical understanding increases. For example:

- including non-medical measures (e.g., lockdown) and environmental factors as inputs to the contagion
- including other vaccines, as well as other variants of the virus

It is clear that the framework can be

- applied to other contagious diseases
- generalized to other observables of interest (e.g., hospitalization and intensive care unit, not discussed here)
- scaled up (e.g., to the European level) or down (e.g., to the level of Communities and provinces)