QUANTIFYING THE IMPACT OF THE KENYAN GOVERNMENT POLICY INTERVENTIONS TO SLOW THE SPREAD OF COVID-19.
(Change Point Analysis.)

INTRODUCTION

The project focuses on quantifying the statistical significance of public health policies introduced by the Kenyan government to slow down the spread of the pandemic, COVID-19.

Data used for the analysis is obtained from the John Hopkins University covid-19 cases database. Precisely, the number of new cases per day will be used to conduct the analysis.

KEY CONCEPTS IN THE ANALYSIS

The purpose of the SIR/SEIR model:

SIR and SEIR are compartmental models that simplify the mathematical modelling of infectious diseases. Subjects progress between the compartments and the order of the labels show the flow pattern. For the SIR model, the compartments are: *Susceptible*, *Infected* and *Recovered*. The SEIR model has an inclusion of the *Exposed* phase (represents subjects who've experienced a long incubation period but are not infectious). Subjects can move from susceptible/exposed to infected compartments through infection by the infected, and from infected to recovery compartments through recovery. The movement between these groups are determined by two main parameters; the *spreading rate* and the *recovery rate.*

Due to the infectious nature of the current pandemic, COVID-19, it can be modelled using these models and perfectly conform to the compartments defined.

The purpose of the models is to describe the dynamics of the expected proportions of subjects in each pandemic state over time using a set of Ordinary Differential Equations. The *infection rates* can also be observed at different points of time to quantify the impact of interventions put in place at different times .

The output of an SIR/SEIR model in relation to COVID-19 cases data:

The overall output of the models is the forecasting of future behaviour of movement of subjects along the compartmental groups through the computation of the *spreading* and *recovery* rates of COVID-19. The prediction will depict different trends at different points of time depending on the impact of the interventions. The output can be used as a reference to provide guidance on which strategies are best fit for the situation and which have no impact.

Processes that affect the generation of the COVID-19 cases data and the parts of these processes modelled by the SIR/SEIR model:

COVID-19 cases data is obtained through mass testing and monitoring of the identified infected subjects.

Processes that affect the generation of the data are taken care of as follows:

- Weekly modulation to accommodate the low cases recorded during weekends that accumulate during the weekdays.

- Inclusion of a reporting delay between new infections and newly reported cases.
- Generalization of simple SIR model with stationary spreading rate for change point detection.

Characteristics of an exponential function:

The graph of an exponential function is asymptotic to the x-axis, passes through the co-ordinates(0,1), is continous and smooth, has a domain set of all real numbers, has a range set of y>0 and can be shifted horizontally or vertically by adding an amount inside or outside the plotting function.

Similarities and differences of an exponential growth, exponential decay, geometric progression, and logistic growth.

An exponential growth is observed when the population of some entities increase rapidly in an exponential manner over time while exponential decay is the exponential decreasing of the values over time.

Logistic growth is observed when the population's growth rate gets smaller with increase in population size, this is imposed by limited environment resources.

A geometric progression is a sequence whereby any element after the first is obtained by multiplying the preceding element by some constant.

The similarity across the 4 is that they all exhibit trends and they involve rapid change in numbers over time.

On the other hand: exponential growth, decay and logistic growth are continous and can be defined for all real numbers while a geometric progression is discrete and is limited only to positive integers.

Approximation that leads the SIR/SEIR model to take an exponential form:

The rate of change at any time, t is dependent on the number of infected and susceptible individuals and the values of the infection and recovery rate parameters.

How the rate parameters in the SIR/SEIR model are estimated from a COVID-19 cases data:

This is achieved through bayesian inferencing using Monte Carlo Markov Chains(MCMC) sampling. MCMC is implemented to obtain the posterior estimates and credible intervals of the unknown parameters. The informative priors are chosen based on available knowledge.

Expected outputs of the modelling phase is:

The spreading rate of COVID-19 at different points of time. This can be represented on a plot to observe changes in the value of the parameter at different points of time.

Predictions to future dates:

Bayesian inferencing on Monte Carlo Marcov Chains samples is done with continous time intergration. Future forecasting is then done based on the results obtained.

How the effectiveness of government COVID19 non-pharmaceutical interventions policies are evaluated :

Change point analysis is done where spreading and recovery rates at each change point are calculated. A bayesian framework is implemented to asses the effects of intervention in a timely manner due to delay between interventions and first useful estimates of the new spreading rates.

ANALYSIS

Why the statistical model is fit for our data

The model used addresses these 3 important steps of crisis mitigation:

- Establishing central epidemiological parameters like the *reproduction number* (the expected number of secondary cases produced by a single infection in a completely susceptible population) for short time forecasting (SIR).
- Simulating effects of interventions put in place to relief the crisis.
- Estimating actual effects of the measures taken not only to make adjustments but to adapt short time forecasts.

Despite the challenges of large statistical and systematic errors that occur during the intitial stages of a pandemic, the model intergrates prior knowledge, assesses the uncertainities about epidemiological parameters quantitatively and principally propagates these uncertainities into focus through the combination of the SIR model with bayesian parameter inferencing (using MCMC sampling) and model augmentation with time dependent spreading rate. The time dependence implementation takes care of the change points in the spreading rate assumed to be driven by governmental interventions and change in individual behaviour, which is the main focus of our analysis: *quantifying the impact of interventions.*

Overview of how the model works and builds on the science publication:

The model performs bayesian inferencing for the central epidemiological parameters of an SIR model using MCMC sampling to compute the posterior distribution of the parameters and to sample from it for forecasting.The SIR (Susceptible-Infected-Recovered) model specifies population compartments and the rate at which they change (susceptible getting infected and the infected recovering). The central parameters are the *spreading rate, recovery rate, the reporting delay* and *the number of initially infected people.* Informative priors are chosen based on available knowledge for the spreading rate, recovery rate and the reporting delay parameters and uninformative priors are chosen for the remaining parameters.

The impact of interventions are quantified by concentrating on the effective growth of active infections before and after the interventions. As long as the population size is greater than the number of infections and recoveries, the number of active infections can be estimated by an exponential growth.

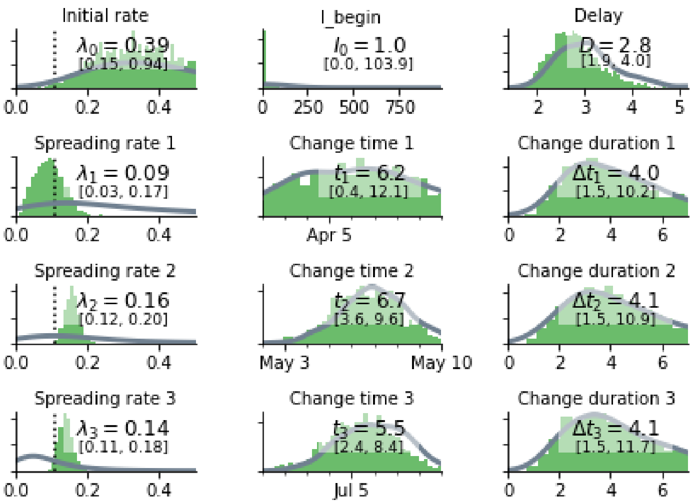
Processes that affect the generation of the data are taken care of as follows:

- Weekly modulation to accommodate the low cases recorded during weekends that accumulate during the weekdays.
- Inclusion of a reporting delay between new infections and newly reported cases.
- Generalization of simple SIR model with stationary spreading rate for change point detection.

The change points and the relative effect of each one and when they were observed:

3 governmental interventions put in place after the cases hit 100 are fit as the change points in our model. Figure 1 below indicates the respective values of *spreading rate* before and after each change point.

Figure 1



- **6th April 2020** – *Ban on movement in Nairobi, Kilifi, Kwale, and Mombasa counties and imposing of the dusk to dawn curfew.*
  The infection rate decreases to 23% (λ_0 = .39 to λ_1 = .09). The assumption is that the 4 counties are the hotspots, thus shutting them down at an early stage (less infected people) greatly reduced the number of contacts between the infected and the susceptible. Although the infection rate declines, the total reported cases continue to grow.
  The date of the change point is inferred to be 5th April, which matches the government's intervention timing.

- **7th May 2020** - *Lockdown in Eastleigh (Nairobi) and Old Town (Mombasa city).*
  The infection rate decreases to 38.5 % (λ_0 = .39 to λ_2 = .16) of the prior which is also an increase compared to change point 1's spreading rate (λ_1 = .09 to λ_2 = .16). At this point, a larger audience has been infected compared to at the time of the first intervention thus as much as the spreading rate is lower than the λ_0 , it cannot be lower than $\lambda_{1.}$ There is a sufficiently large number of contacts between the infected and the susceptible.
  The date of the change point is inferred to be around 7th May, which matches the government's intervention timing.

- **6th July 2020** - *Movement restrictions lifted in Nairobi, Mombasa, and Mandera counties. ('One man for himself': encouraging use of masks and everyone to take care of themselves)*
  The spreading rate is way lower than the prior (λ_0 = .39 to λ_3 = .14) , and skightly lower than $\lambda_{2.}$ (λ_2 = .16 to λ_3 = .14) .The expectation of lifting of movement restrictions is an increase in contact thus an increase in the spreading rate. The slight reduction from λ_2 has been observed instead. This can be explained by the non-pharmaceutical intervention of encouraging every person to take care of himslef. As much as the contacts increase, people are extra cautious thus its a fair trade.
  The date of the change point is inferred to be 5th July, which matches the government's intervention timing.

The predicted number of cases for the country for the week of 10 August 2020:

	date	predicted new_cases
0	2020-08-08	740
1	2020-08-09	652
2	2020-08-10	532
3	2020-08-11	583
4	2020-08-12	692
5	2020-08-13	761
6	2020-08-14	777

The predictions indicate reduction in the new cases for the first few days then a rise.

Limitations of the analysis:

- The model doesn't fit the recovered cases and yet the *recovery rate* parameter is essential in the forecasting.
- Alot of model assumptions are based on the situation in Germany. This might not be the case for Kenya.

Lessons learnt that countries could adopt for future pandemics:

- Mitigation policies should be implemented during the intial stages of a pandemic. It is when the situation can be put under control easily.
- Policies that foster reduction of contact between people should always be considered.
- General personal care should always be advocated for throughout, not only during pandemics. When this becomes normal among people, it brings major impact in controlling the spread of infectious diseases.

CONCLUSION

Different interventions have different degrees of impact towards slowing the spread of COVID-19. Knowing the level at which a particular intervention is significant helps in steering the country in the right direction both in the current situation and in the future.

REFERENCES

1. Jonas Dehning, Johannes Zierenberg et.al, 'Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions', *Science* 10 Jul 2020: Vol.369, Issue: 6500eabb9789, DOI: 10.1126/science.abb9789 .
2. https://github.com/Priesemann-Group/covid19_inference_forecast/
3. https://jrnold.github.io/bayesian_notes
4. Scientific Brief, 'Estimating mortality from COVID-19'(4 August 2020).