

TimeLocal Reproduction Number

Counting Processes, Martingales,
Log-Likelihood MLE's and Variances

TimeLocal Reproduction Number

Counting Processes, Martingales,
Log-Likelihood MLE's and Variances

Jan Donker

Writer: Jan Donker
Coverdesign: Jan Donker
ISBN: 9789464927375
© 2024 by Jan Donker

Preface

About a 100 years ago, biochemist William Ogilvy Kermack and epidemiologist Lieutenant Colonel Anderson Gray McKendrick launched one of the first mathematical models in epidemiology in the form of a system of differential equations, the SIR model, dealing with the different parts of the population when a pandemic is at hand: 'S' for 'Susceptible', 'I' for 'Infectious or Infected', 'R' for 'Recovered or Removed', and how individuals travel from each of these 'compartments' S,I,R to another. There have been several extensions since then, but nevertheless, the original SIR model has been the most studied, not only deterministically, but of late stochastically and also numerically.

The main parameters used and estimated are the transmission rate λ , duration of infectious period $1/\gamma$, and the ensuing basic reproduction number $R_0 = \lambda/\gamma$, but, as one average value over the whole duration of the epidemic, noted as $[0, T]$, T the end of the pandemic. Another version of a time-varying reproduction number is a weighted average based on the Euler-Lotka equation, where still the period $[0, t]$, $t \leq T$, is the time-varying factor in stead of a point t in time itself.

This seems to be the norm in really all the investigations I at least have seen.

In retrospect eying the apparent uncertainties of the different governments all over the world with respect to handling the COVID-19 pandemic in the midst of crisis, it's no wonder that the daily or weekly reported increases of numbers of new infections, certainly in the public eye, took precedence over a simultaneously even slightly reclining reproduction number as happened here in the Netherlands.

Monitoring the course of a pandemic by treating parameters like the transmission rate and the duration of the infectious period as functions of time, the basic timelocal reproduction number, for now defined as $R_t^{\text{bloc}} = \lambda(t)/\gamma(t)$, while setting $R_0 = \lambda(0)/\gamma(0)$, more adequately reflects the direct effects of real changes of government policy during the pandemic, like social distancing, vaccination, disease treatment, medicin use, or for that matter, any significant change to the course of the pandemic. I would not exclude wars or alien visitors from outer space as possible gamechangers!

This work is about a stochastic approach to the problem. For the SIR model, at the turn of the millennium, Britton and Anderson had worked out a weighted averaging method for a period $[0, t]$. What they constructed is an algorithm that has as it's input 0 as the start of the period and t as a moving end of the period. Simply enough, my contribution has been to change the period to $[t, t + h]$, locally around time t , thereby obtaining a moving average. At the risk of selling myself short, a simple Taylor expansion of the function $\lambda(t)/\gamma(t)$ around t in terms of the SIR system of differential equations confirmed the relevance of my sumptions: It is about the local growth rate in time, the daily/weekly reported increases/decreases of numbers of infections. From there on it was helpful that the stochastic processes to be studied were Markovian.

I did however, apart from working it out for the SIR model, expand on their method to a SEIR with a birth rate and death rate. This is preceded by first formulating and giving context to the timelocal reproduction number as essentially a product of several 'single transmission and transition reproduction numbers', in case there are 1 or more latency states: 'E' for 'Exposed: infected but not yet infectious'. Clearly structuring, explaining and clarifying the aforementioned algorithm at hand, sustained by completely written out proofs, is a large part of the effort also, since for most of us simple earthly dwellers, mathematical statistics is not exactly an easy walk in the park.

Jan Donker

Contents

| | |
|---|-----------|
| Introduction Timelocal Reproduction Number (TLRN) | 1 |
| 1 From Counting Processes to Martingales for TLRN | 9 |
| 1.1 Counting Processes Generating Centered Poisson Processes as Zero Mean Martingales | 9 |
| 1.2 Mutually Independent Zero Mean Martingales: No Simultaneous Jumps | 19 |
| 2 Back to Differential Equations by Way of Integral Equations | 22 |
| 2.1 Intensities $\lambda_1(t + s)$ and $\lambda_2(t + s)$ Timelocal SIR Model Markovian Version | 22 |
| 2.2 Almost Sure Convergence of $\frac{Y_i^t(s)}{n}$ for $s \in [0, h]$ to Deterministic Solutions $y_i^t(s)$ of Integral Equations | 24 |
| 2.3 Applying Vitali's Convergence Theorem | 26 |
| 3 Variation Processes for Martingales $M_{i,f_i}^t(h) = \int_0^h f_i(s)(dY_i^t(s) - \lambda_i(t + s)ds)$ $i = 1, 2$ | 28 |
| 3.1 Definitions..... | 28 |
| 3.2 Calculating Variances and Expectations | 28 |
| 3.3 Applying Rebolledo's Theorem for Martingales | 33 |
| 3.4 Consequences of Rebolledo's Theorem and Interaction with MLE | 34 |
| 4 Log-Likelihood Functions and MLE's $\hat{\lambda}_{t,h}, \hat{\gamma}_{t,h}$ for True Values $\lambda_{t,h}$ and $\gamma_{t,h}$ | 37 |
| 4.1 Shape of Time-Dependent Distribution and Preview to MLE | 37 |
| 4.2 Determining the Log-Likelihood Functions and MLE's | 39 |
| 4.3 Calculating the Log-Likelihood for the Distribution of $Y^t(h)$ | 41 |
| 4.3.1 Contemplating the Pmf on Microlevel | 44 |
| 5 Calculating the Means and Variances for $\hat{\lambda}_{t,h}, \hat{\gamma}_{t,h}$ and $\hat{R}_t^{bloc}(h) = \frac{\hat{\lambda}_{t,h}}{\hat{\gamma}_{t,h}}$ | 45 |
| 5.1 For $\hat{\lambda}_{t,h}$ and $\hat{\gamma}_{t,h}$ | 45 |
| 5.2 For $\hat{R}_t^{bloc}(h) = \frac{\hat{\lambda}_{t,h}}{\hat{\gamma}_{t,h}}$ | 48 |
| 6 Analysis Preferred Estimate of Effective Reproduction Number: | |
| $\hat{R}_t(h) = \frac{S(t)}{n} \hat{R}_t^{bloc}(h)$ or $\hat{R}_t(h) = \frac{\int_0^h \bar{S}(t+s)I(t+s)ds}{\int_0^h I(t+s)ds} \hat{R}_t^{bloc}(h) = \frac{Y_1^t(h)}{Y_2^t(h)}$? | 52 |
| 6.1 Exponential Distribution Time Variable for $\hat{R}_t(h) = \frac{Y_1^t(h)}{Y_2^t(h)}$? | 52 |
| 6.2 For $\hat{R}_t(h) = \frac{S(t)}{n} \hat{R}_t^{bloc}(h) \vee \frac{Y_1^t(h)}{Y_2^t(h)}$ | 54 |
| 6.2.1 Making the Case: Differences between $\hat{R}_t(h) = \frac{S(t)}{n} \hat{R}_t^{bloc}(h)$ and $\hat{R}_t(h) = \frac{Y_1^t(h)}{Y_2^t(h)}$ | 54 |
| 6.2.2 Determination of Expectation and Variance of $\hat{R}_t(h) = \frac{Y_1^t(h)}{Y_2^t(h)}$ (Version 2) | 55 |
| 6.2.3 Determination of Expectation and Variance of $\hat{R}_t(h) = \frac{S(t)}{n} \hat{R}_t^{bloc}(h)$ (Version 1) | 57 |
| 6.3 Conclusion | 59 |

| | |
|---|------------|
| 7 Extension of SIR to the SEIR model with Birth Rate and Death Rate..... | 60 |
| 7.1 The Effective Reproduction Number for more than One State of Infection..... | 60 |
| 7.2 ‘Homogeneous’ Extension of SIR to SEIR: Adding Latency Period..... | 75 |
| 7.3 ‘Inhomogeneous’ Extension of SIR to SEIR: Adding Birth Rate and Death Rate | 77 |
| 8 Linking the Effective Reproduction Number (R_t) of the Euler-Lotka Population Growth Model used by RIVM with the Timelocal RN | 91 |
| 8.1 Crunching the Numbers to Obtain a Discrete Formulation of R_t from Scratch | 91 |
| 8.2 Comparison Continuous Kermack-McKendrick with Discrete Euler-Lotka and Short Statistical Formulation | 94 |
| 8.3 Critique of $RN_{RIVM}(k)$ | 95 |
| 8.4 Further Examples of Behaviour $RN_{RIVM}(k)$ | 97 |
| 9 Appendix | 101 |
| 9.1 Approximations for Mean and Variance of a Ratio | 101 |
| 9.2 Convergence in Probability \rightarrow Convergence in $\mathcal{L}^p, p \geq 1$ with Additional Condition: | 103 |
| Vitali’s Convergence Theorem | |
| 9.3 Gaussian Weak Convergence in Distribution of Centered Poisson Processes | 105 |
| (Donsker’s Central Limit Convergence Theorem) | |
| 9.4 Different Definitions of the Effective Reproduction Number | 109 |
| 9.5 How to Locally Linearize to Kermack-McKendrick Equations with a Birth Rate | 111 |
| 9.6 Original Inspiration: SIR Model for a Timelocal Reproduction Number (TLRN) | 114 |
| 10 Afterword: Evaluation | 115 |
| 11 References | 119 |
| 12 Secondary References | 121 |

Introduction Timelocal Reproduction Number (TLRN)

To understand what is meant by a timelocal reproduction number, we have to establish a clear link with the original concept of a reproduction number, starting with the original deterministic Kermack-Mckendrick equations, where the following parameters are used to shape the reproduction number:

- λ (Transmission Rate) =

C (Number of Contacts per Unit of Time) . P (Probability of Transmission per Contact)

- $\frac{1}{\gamma} = D$ (Duration of Contagious, Infectious Period)

$S'(t) = -\frac{\lambda}{N}S(t)I(t)$ together forming the reproduction number $R_0 = \frac{CPD \cdot S(0)}{N} = \frac{\lambda S(0)}{N\gamma}$, which decides if

$I'(t) = \frac{\lambda}{N}S(t)I(t) - \gamma I(t)$ the epidemic takes off: $R_0 > 1 \rightarrow I'(0) > 0$, or doesn't start: $R_0 < 1 \rightarrow I'(0) < 0$.

$R'(t) = \gamma I(t)$ Usually λ and γ , and thus R_0 , are taken as constant, and when the epidemic is over,

$S(t) + I(t) + R(t) = N$ the part of the population infected, z , is determined by: $1 - z = e^{-R_0 z}$.

$N = n + m$: total population In this situation, since $S(0) = n$: $R_0 = \frac{\lambda S(0)}{N\gamma} = \frac{n\lambda}{N\gamma} = \frac{n}{n+m} \frac{\lambda}{\gamma} \approx \frac{\lambda}{\gamma}$ since $m \ll n$

$S(0) = n, I(0) = m, R(0) = 0, m \ll N$

$S(t) =$ Susceptible,

$I(t) =$ Infected and Infectious

$R(t) =$ Recovered or Removed

Of course, the actual rate at which $I(t)$ grows or shrinks is $\frac{S(t)}{N}\lambda - \gamma$ resulting in a factual

time-dependent reproduction number $R_{eff}(t) = \frac{\lambda S(t)}{N\gamma} = R_0 \frac{S(t)}{S(0)} = R_0 \frac{S(t)}{n}$ and as functions of time:

$$I(t) = I(0)e^{\int_0^t (\frac{\lambda}{N}S(s) - \gamma) ds} = I(0)e^{-\gamma t + \int_0^t \frac{\lambda}{N}S(s) ds} \quad (\text{Taking } I(0) > 0)$$

$$R(t) = \int_0^t \gamma I(u) e^{-\gamma u + \int_0^u \frac{\lambda}{N}S(s) ds} du \quad \text{so exponential growth until } \int_0^t \frac{\lambda}{N}S(s) ds \leq \gamma t \text{ then growth is dying out.}$$

The factor $\frac{S(t)}{n}$, with $S(t)$ a declining function of time, secures the existence of a 'herd immunity' or

rather 'population immunity', which results in the already referred to formula: $1 - z = e^{-R_0 z}$,

describing the part of the total population infected after the end of the pandemic ($I'(t) = I(t) = 0$).

That's all very fine, but what if you want to take into consideration a curbed contact rate, ensconced in λ , by social distancing as a government regulation, or a curbed duration of the infectious period, represented by $\frac{1}{\gamma}$, by using Remdesivir or Regeneron, like Donald Trump?

Mr. President 'The Donald' probably thought he had been curbing the epidemic singlehandedly on his own, actually claiming the pandemic had all but subsided, while it was raging as never before.

A true example of 'Trump Statistics', also known as 'Foney Math'! Ooh I'm sorry, 'Phoney Math' or 'Fake Math'. I'm told he just lost reelection but couldn't and wouldn't believe it. Go Joe!

Joking aside, for now you would want to 'mathematically' (actually: statistically!) prove him wrong on the spot, but alas, so far you're stuck! Don't you hate it when that happens?

Well, as a solution, you could want to take λ and γ as functions of time $\lambda(t)$ and $\gamma(t)$ to get to:

$S'(t) = -\frac{\lambda(t)}{n}S(t)I(t)$ forming the effective (timelocal) reproduction number

$I'(t) = \frac{\lambda(t)}{n}S(t)I(t) - \gamma(t)I(t)$ $R^{loc}(t) = \frac{S(t)\lambda(t)}{n\gamma(t)} = R_{eff}(t)$, writing $R_t^{bloc} = \frac{\lambda(t)}{\gamma(t)}$ as the

$R'(t) = \gamma(t)I(t)$ basic timelocal reproduction number where still the epidemic takes off at $t = 0$ or not, the same way:

$$S(t) + I(t) + R(t) = N = n + m \quad R_0 = R^{loc}(0) = R_{eff}(0) = \frac{\lambda(0)S(0)}{n\gamma(0)} = \frac{\lambda(0)}{\gamma(0)} = R_0^{bloc}.$$

$N = n + m$: total population

and as functions of time:

$$I(t) = I(0)e^{\int_0^t (\frac{\lambda(s)}{n}S(s) - \gamma(s))ds},$$

$$R(t) = \int_0^t \gamma(u)I(0)e^{\int_0^u (\frac{\lambda(s)}{n}S(s) - \gamma(s))ds} du, \text{ so exponential growth until } \int_0^t \frac{\lambda(s)}{n}S(s)ds \leq \int_0^t \gamma(s) ds,$$

then growth dying out.

Again, the factor $\frac{S(t)}{n}$, with $S(t)$ a declining function of time, here not so much securing the existence of a 'population immunity', but surely backing up a successful end to the epidemic (Are there still going to be enough of us to go on?), if parameters $\lambda(t)$ and $\gamma(t)$ allow so. They probably don't if you hug too much!

Silently we have changed the expression $\frac{\lambda}{N}S(t)I(t)$ to $\frac{\lambda}{n}S(t)I(t)$ (here $\frac{\lambda(t)}{n}S(t)I(t)$). This need not necessarily change the overall equation since λ and γ or $\lambda(t)$ and $\gamma(t)$ for now are unknown anyway, and so you can simply assume for the 'original' deterministic equation: $R_{0,N} = \frac{n}{N} \frac{\lambda_N}{\gamma_N} = \frac{\lambda_n}{\gamma_n} = R_{0,n}$.

So basically, numerically $\lambda_N \approx \lambda_n$, $\gamma_N = \gamma_n$, and n can function as population size as easily as N does. If for instance $m = 1$ and $N = 20.000.000$, of course the difference is ridiculously small!

The same for $\lambda(t)$ and $\gamma(t)$. And so this can be interpreted more or less as a simple rescaling.

This all comes from the concept of the so-called Sellke construction which gives rise to the stochastic intensity process $\frac{\lambda}{n}S(t)I(t)$, where $n = S(0)$, as the rate by which new individuals become infected.

The Sellke construction investigates how the cumulative force of infection $A(t)$ works: $A(t) = \frac{\lambda \int_0^t I(s)ds}{n}$. So $A(t)$ increases at rate $\frac{\lambda I(t)}{n}$ ($= A'(t)$?). In par. 3.5.2, page 64 of [1], it is shown that the Sellke construction accurately describes the definition or set up of the 'prototype stochastic epidemic model' (par. 3.1 [1]), of which we apply, as far as the 'standard' Kermack-McKendrick equations are concerned, the special case of the 'general' stochastic epidemic (par. 3.2.2 of [1]), for which no latency period is assumed, and which is endowed with the Markov property. The Markov property is there because the infectious period is assumed to be exponentially distributed, so memoryless, hence Markovian. This implies that each susceptible individual becomes infected at rate $\frac{\lambda I(t)}{n}$.

Because there are $S(t)$ such individuals at any time t the overall rate at infection becomes $\frac{\lambda}{n}S(t)I(t)$.

Par. 5.4 [1] and chapters 2,3, 4, 5 and 9 [2] deal extensively with this concept in the Kermack-McKendrick case. The Sellke construction does contain the possibility of working with a latency period. However, when studying an extended SEIR example in chapter 7, in par 7.3, which includes an in time variable stochastic population size, $N(t)$, we will revert to $n = \frac{\int_0^h N(t+s)ds}{h}$ as the average population size on $[t, t + h]$, since, as we have seen, the two situations differ very little.

At the end of paragraph 7.3 a short investigation of the difference between the two is done.

We now have seen that $R^{loc}(t) = \frac{S(t)}{n} \frac{\lambda(t)}{\gamma(t)} = R_{eff}(t)$ for both $\frac{\lambda(t)}{\gamma(t)}$ and $\frac{\lambda}{\gamma}$ constant in time.

Why we still differ between the effective reproduction number and the timelocal reproduction number, is because in other situations like Euler-Lotka (see for instance [14] or [6]), they are not the same. There, $R_{eff}(t)$ is really a function of the interval $[0, t]$, or, in the discrete case, of the array $[1, 2, \dots, k]$, so 'effectively', though in fact only optically, a function of t or k .

Posing the Kermack-McKendrick problem discretely, time steps are actually generation steps. Switching from time t to generation k , causes $\gamma(k) = 1$, and $\lambda(k)$ ensconced in $I(k + 1)$:

$$I_{k+1} - I_k = \left(\frac{S(k)}{n} \lambda(k) - \gamma(k) \right) I_k \xrightarrow{\gamma(k)=1} I_{k+1} = \frac{S(k)}{n} \lambda(k) I_k \rightarrow R^{loc}(k) = \frac{S(k)}{n} \lambda(k) = \frac{I_{k+1}}{I_k}$$

So in fact, in this case, we are really looking at only a generation-varying infection rate, which looks like a time-varying infection rate as mentioned in [6], but isn't. Since in [6], the term 'time-varying' refers more to a time-varying interval $[0, t]$. Therefore I specifically have used the term 'timelocal'. For a further clarification on the subject see also appendix 9.4. This is preceded in chapter 8, par. 1 and par. 2, by an analysis of the characteristics of the effective reproduction number, as it is defined in [6] and [14], and allegedly used by the RIVM in the Netherlands during the COVID-19 pandemic.

The RIVM uses the Euler-Lotka population growth model as described in article [6] (Wallinga and Lipsitch), at least judging from their website announcement at the time, where this is developed into, it seems to me, a central formula for R_t :

$$R_t = \frac{\hat{b}(t)}{\int_0^\infty \hat{b}(t-a)g(a)da} \quad (\text{or: } \hat{b}(t) = R_t \int_0^\infty \hat{b}(t-a)g(a)da)$$

Van Dissel of the RIVM talks about R_t as 'the number of secondary infections per case' and in [6] it is: 'This reproductive number R_t assigns it value to the time at which the secondary cases are infected'. So this seems to be a strong indication that we have the right formula. The above formula is looking a lot like the formula Tom Britton sent to me, which is: (compare to: $\hat{I}(t) = R_t \int_0^\infty \hat{I}(t-a)g(a)da$)

$$I(t) \sim \text{Poisson}(R_t \sum_k I(t-k)g(k))$$

I found some flaws with the predictive value of this reproduction number, when daily reported infection rates, like for instance $\frac{I_{k+1}}{I_k}$, remain high. Paragraphs 8.3 and 8.4 deal with this situation.

You could call the expression $\frac{I_{k+1}}{I_k}$ the 'generationlocal reproduction number', if you like.

The MLE for statistically describing the basic reproduction number R_0 after the end of the pandemic ($I'(T) = I(T) = 0$), is noted $\hat{R}_0(T)$ where $\hat{R}_0(t)$ is calculated by the formula: ([1] and [2] just use \hat{R}_0)

$$\hat{R}_0(t) = \frac{\hat{\lambda}}{\hat{\gamma}} = \frac{N_1(t) - N_1(0)}{\int_0^t I(s)\bar{s}(s)ds} \frac{\int_0^t I(s)ds}{N_2(t) - N_2(0)}, \quad N_1(t) = I(t) + R(t) \text{ and } N_2(t) = R(t) \text{ to be found in [1] and [2].}$$

This looks like a function of time, but really isn't. Here, $\hat{R}_0(t)$ is just a one value estimate for the whole period $[0, t]$, and moreover, $\hat{R}_0(t) \geq 1$ always holds, no matter what, which is logical anyway, since $\hat{R}_0(t)$ represents R_0 in case the epidemic actually takes off. [1] and [2] just use the expression \hat{R}_0 , but I use the expression $\hat{R}_0(t)$, because soon, the number of parameters like this will become more than 1. In fact, as a function of time, $\hat{R}_0(t)$ describes the situation of the pandemic as a whole, averaging out intermediate spikes in $\frac{\lambda(t)}{\gamma(t)}$. The part of the total population infected is $\frac{R(T)}{N}$.

As an extra later on, in chapter 8, we will derive an alternative formula for $R(t)$ and $R(T)$ in terms of the timelocal reproduction number in a discretely posed problem, regarding the effective reproduction number, as it is defined in [6]. For the original inspiration which led to the concept of a TLRN, see appendix 9.6.

Furthermore, the actual timelocal aspect of the effective/timelocal reproduction number is in fact the expression $\frac{\lambda(t)}{\gamma(t)}$ rather than $\frac{S(t)}{n}$. For instance, social distancing primarily affects $\lambda(t)$.

This partly explains the term 'basic' in the basic timelocal reproduction number $R_t^{bloc} = \frac{\lambda(t)}{\gamma(t)}$.

Of course, here, 'basic' doesn't mean 'epidemic taking off or not' as the 0 in R_0 means, but, as the basic reproduction number R_0 is estimated as one value by $\hat{R}_0(t)$ for the interval $[0, t]$, we estimate the basic timelocal reproduction number R_t^{bloc} as one value for the interval $[t, t+h]$ by $\hat{R}_t^{bloc}(h)$.

So there is never a $R_0(t)$, but only a statistical estimate $\hat{R}_0(t)$ for R_0 on the interval $[0, t]$, and also there is never a $R_t^{bloc}(h)$, but only a statistical estimate $\hat{R}_t^{bloc}(h)$ for R_t^{bloc} on the interval $[t, t+h]$. This explains why we always have $\hat{R}_0(t) > 1$ (epidemic taken off!) but for $\hat{R}_t^{bloc}(h)$ we only have $\hat{R}_t^{bloc}(h) \geq 0$. This will be discussed extensively later on in this introduction.

As statistical estimates for $R_{eff}(t) = R^{loc}(t) = \frac{S(t)}{n} R_t^{bloc}$ on interval $[t, t+h]$ we will use $\hat{R}_t(h)$ for $R_{eff}(t)$ and $\hat{R}_t^{loc}(h)$ for $R^{loc}(t)$. Of course for now we have: $\hat{R}_t(h) = \hat{R}_t^{loc}(h)$.

The factual reproduction number might be $R_{eff}(t) = R^{loc}(t) = \frac{S(t)}{n} R_t^{bloc}$, $\frac{S(t)}{n}$ as prefactor is just a multiplication number which always applies. We are mainly interested in statistically determining $\frac{\lambda(t)}{\gamma(t)}$.

Regarding COVID-19, the first year sees this multiplication number $\frac{S(t)}{S(0)}$ probably mainly between 0,97 and 1, in fact mostly between 0,99 and 1. This factor seems to be an object for data research only later in the epidemic. Chapter 6, amongst others, deals with this. Now how does R_t^{bloc} look in terms of the above differential equations? Like this: ($R^{loc}(t)$ looks even less complicated)

$$R_t^{bloc} = \frac{\lambda(t)}{\gamma(t)} = \frac{n}{S(t)} \frac{\lambda(t)S(t)I(t)}{n} \frac{1}{\gamma(t)I(t)} = \frac{n}{S(t)} \frac{I'(t)+R'(t)}{R'(t)} = \frac{n}{S(t)} \left(1 + \frac{I'(t)}{R'(t)}\right) \quad (R^{loc}(t) = \frac{S(t)}{n} \frac{\lambda(t)}{\gamma(t)} = 1 + \frac{I'(t)}{R'(t)})$$

Derivatives relating to jump processes are difficult to administer statistically, but we can use the difference quotient to reveal the discrete random variables related to the underlying counting processes. For h small enough, we use Taylor expansion $f(t+h) - f(t) = hf'(t) + o(h)$ where here mainly is used:

$$\text{Let } f_{a,b}(t) = \frac{a-t}{b-t}, a > 0 \text{ and } b > 0, \text{ then } f_{a,b}(o(h)) = \frac{a-o(h)}{b-o(h)} = \frac{a}{b} + \frac{a-b}{b^2} o(h) + o(h) = \frac{a}{b} \pm o(h).$$

In the formulas below, in the first step, we in fact use: $hf'(t) = f(t+h) - f(t) - o(h)$ for numerator and denominator, and in the second step the form $f_{a,b}(t)$ for the whole expression.

$$R_t^{bloc} = \frac{\lambda(t)}{\gamma(t)} = \frac{n}{S(t)} \frac{h(I'(t)+R'(t))}{hR'(t)} = \frac{n}{S(t)} \frac{I(t+h)-I(t)+R(t+h)-R(t)-o(h)}{R(t+h)-R(t)-o(h)} = \frac{n}{S(t)} \frac{N_1(t+h)-N_1(t)}{N_2(t+h)-N_2(t)} \pm o(h)$$

$$(R^{loc}(t) = \frac{S(t)\lambda(t)}{n\gamma(t)} = \frac{N_1(t+h)-N_1(t)}{N_2(t+h)-N_2(t)} \pm o(h) = \frac{I(t+h)-I(t)+R(t+h)-R(t)}{R(t+h)-R(t)} \pm o(h) = 1 + \frac{I(t+h)-I(t)}{R(t+h)-R(t)} \pm o(h))$$

with $N_1(t) = I(t) + R(t)$ and $N_2(t) = R(t)$ as the stochastic processes/random variables here.

$$\text{Expanding } \frac{n}{S(t)} \text{ over the interval } [0, h]: \frac{n}{S(t)} = \frac{hnI(t)}{hS(t)I(t)} = \frac{n \int_0^h I(t+s)ds}{\int_0^h S(t+s)I(t+s)ds} \pm o(h)$$

We move to an estimate for R_t^{bloc} : (using $(a \pm o(h))(b \pm o(h)) = ab \pm o(h)$)

$$R_t^{bloc} = \left(\frac{N_1(t+h)-N_1(t)}{N_2(t+h)-N_2(t)} \pm o(h) \right) \left(\frac{n \int_0^h I(t+s)ds}{\int_0^h S(t+s)I(t+s)ds} \pm o(h) \right)$$

$$= \frac{N_1(t+h) - N_1(t)}{\int_0^h I(t+s)S(t+s)ds} \frac{\int_0^h I(t+s)ds}{N_2(t+h) - N_2(t)} \pm o(h) = \hat{R}_t^{bloc}(h) \pm o(h)$$

This might look like we have been blurring the estimate, but this will turn out to be the MLE resulting from the log-likelihood equation, that comes out of the counting processes and corresponding jump processes we are going to describe, as it is mentioned on the previous page as well. Furthermore, an MLE, in this case, should be spread naturally over the interval $[0, h]$. Finally, showing the soundness of this estimate, by taking the limit for $h \rightarrow 0$:

$$\begin{aligned} \lim_{h \rightarrow 0} \hat{R}_t^{bloc}(h) &= \lim_{h \rightarrow 0} \frac{\frac{N_1(t+h) - N_1(t)}{h} \frac{n \int_0^h I(t+s) ds}{h}}{\frac{\int_0^h I(t+s) S(t+s) ds}{h} \frac{N_2(t+h) - N_2(t)}{h}} = \frac{N_1'(t) I(t)}{I(t) \bar{S}(t) N_2'(t)} = \frac{N_1'(t)}{N_2'(t) \bar{S}(t)} = \frac{I'(t) + R'(t)}{R'(t) \bar{S}(t)} \\ &= \frac{n}{S(t)} \left(1 + \frac{I'(t)}{R'(t)}\right) = \frac{\lambda(t)}{\gamma(t)} = R_t^{bloc} \cong 1 + \frac{I'(t)}{R'(t)} \text{ in the first few months of the epidemic.} \end{aligned}$$

The formula for $\hat{R}_t^{bloc}(h)$ is in sync with formula 5.41, par. 5.4.2 of [1] for $\hat{R}_0(t)$, in the sense that only the observed timeperiod is different: interval $[0, t]$ replaced by interval $[t, t + h]$.

$$\hat{R}_0(t) = \frac{N_1(t) - N_1(0)}{\int_0^t I(s) \bar{S}(s) ds} \frac{\int_0^t I(s) ds}{N_2(t) - N_2(0)} \approx \frac{N_1(t) - N_1(0)}{N_2(t) - N_2(0)} = 1 + \frac{I(t) - I(0) + R(t) - R(0)}{R(t) - R(0)} \cong 1 + \frac{I(t)}{R(t)}$$

in the first few months of the epidemic.

Since $I(t)$, by it's very nature, in fact is more of a timelocal number, so to speak the leaky tire you are constantly inflating, and which constantly leaks to (feeds) $R(t)$, it is no riddle that this reproduction number stays modest, however large $I(t)$, locally in time, gets. Even if you take into account the

$$\text{multiplication factor } \frac{\int_0^t I(s) ds}{\int_0^t I(s) \bar{S}(s) ds} = \frac{n \int_0^t I(s) ds}{\int_0^t I(s) S(s) ds} \approx 1 \text{ which in practice will stay like that for a long time.}$$

Of course, $\hat{R}_0(t)$ does reflect changes like government measures (affecting $N_1(t), I(t), S(t)$), changes in the infectious period (affecting $N_2(t), I(t), S(t)$), and so on, but does so in an aggregated way, on average if you will. More or less containing $R(t)$, making it grow less or more.

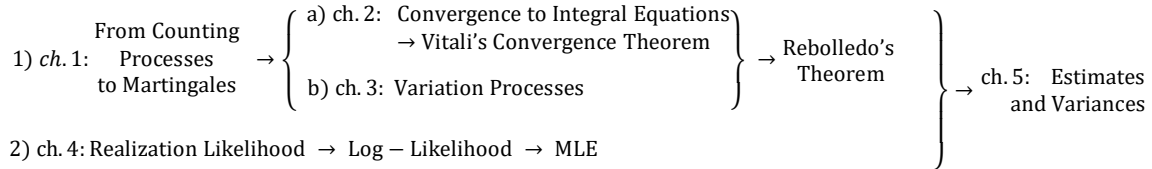
As far as I can see, there are two ways to estimate $R_{eff}(t)$:

$$1) \hat{R}_t(h) = \frac{S(t)}{n} \hat{R}_t^{bloc}(h) = \frac{S(t)}{n} \frac{N_1(t+h) - N_1(t)}{\int_0^h I(t+s) \bar{S}(t+s) ds} \frac{\int_0^h I(t+s) ds}{N_2(t+h) - N_2(t)}$$

$$2) \text{ Expanding } \frac{S(t)}{n} \text{ over the interval } [0, h] \text{ like the expression } \frac{n}{S(t)} \text{ before: } \frac{S(t)}{n} = \frac{\int_0^h \bar{S}(t+s) I(t+s) ds}{\int_0^h I(t+s) ds} \pm o(h):$$

$$\hat{R}_t(h) = \frac{\int_0^h \bar{S}(t+s) I(t+s) ds}{\int_0^h I(t+s) ds} \hat{R}_t^{bloc}(h) = \frac{N_1(t+h) - N_1(t)}{N_2(t+h) - N_2(t)} = 1 + \frac{I(t+h) - I(t)}{R(t+h) - R(t)} \left(\xrightarrow{h \rightarrow 0} 1 + \frac{I'(t)}{R'(t)} \right)$$

My preference is nr. 2, which shouldn't be surprising. It is important to note that $I(t + h) - I(t) < 0$ is a possibility which then means that $\hat{R}_t(h) < 1$. All of this will be covered extensively further on in chapter 6. The main part of this document happens in chapters 1 – 5:



1) Ch.1: The first path formulates stochastic counting processes $Y_i^t(h) = N_i(t + h) - N_i(t)$ on basis of Kermack-McKendrick equations ($N_1(t) = I(t) + R(t), N_2(t) = R(t)$), notation of fractions: $\frac{Y_i^t(s)}{n}$, with

$$\text{matching jump processes } \begin{cases} P(Y_i^t(h + \delta) - Y_i^t(h) = 1) = \delta \lambda_i(t + h) + o(\delta) \\ P(Y_i^t(h + \delta) - Y_i^t(h) = 0) = 1 - \delta \lambda_i(t + h) + o(\delta) \\ P(|Y_i^t(h + \delta) - Y_i^t(h)| = 0) = 1 - \delta \sum_{i=1}^2 \lambda_i(t + h) + o(\delta) \end{cases} \begin{array}{l} \text{with intensities } \lambda_i(t + s), i = 1, 2 \\ \text{by formulating time - dependent} \\ \text{probability distributions,} \end{array}$$

resulting in formulating corresponding zero mean martingales: $M_i^t(h) = Y_i^t(h) - \int_0^h \lambda_i(t + s) ds$,

with matching intensities on $[t, t + h]$: $\lambda_1(t + s) = \frac{\lambda}{n} S(t + s -) I(t + s -)$ and $\lambda_2(t + s) = \gamma I(t + s -)$.

These are connected with, part of, the original martingales in [1] and [2] for the whole interval $[0, t]$:

$M_i(t) = N_i(t) - N_i(0) - \int_0^t \lambda_i(s) ds$ where $\lambda_1(s) = \frac{\lambda}{n} S(s-)I(s-)$ and $\lambda_2(s) = \gamma I(s-)$. As we have

seen, the $M_i(t)$ -parameters λ and γ are to be estimated by $\hat{R}_0(t) = \frac{\hat{\lambda}}{\hat{\gamma}} = \frac{N_1(t) - N_1(0)}{\int_0^t I(s) S(s) ds} \frac{\int_0^t I(s) ds}{N_2(t) - N_2(0)}$

depending on the interval $[0, t]$.

In themselves, the $M_i^t(h)$ do not represent a mini epidemic in the same sense that the $M_i(t)$ represent an epidemic in progress, because this would require $n = n(t) = S(t)$ for the beginning.

It is more like cutting a ' $[t, t+h]$ - slice' of the ' $[0, t]$ - martingale', and then forming estimates for λ and γ , while taking into account that the actual 'slice' is represented by the interval $[t, t+h]$.

It's all Markovian anyway, and this is why the time-dependent probability distributions connected to the jump process are not explicitly conditionally formulated. This approach is consistent with the preliminary MLE estimate we found:

$$\hat{R}_t^{bloc}(h) = \frac{N_1(t+h) - N_1(t)}{\int_0^h I(t+s) S(t+s) ds} \frac{\int_0^h I(t+s) ds}{N_2(t+h) - N_2(t)}$$

I hope my 'chunk of pie wise' explanation didn't put you off too much! I'm fond of pie, sometimes... Consequently the $M_i^t(h)$ are identified as centered Poisson processes, thereby ensuring two things:

a) Ch. 2: Integral equations as deterministic limits of almost sure convergence of random variables:

When the $Y_i^t(s), i = 1,2$ are written as fractions $(\frac{Y_i^t(s)}{n}, i = 1,2), s \in [0, h]$, then for $n \rightarrow \infty$, we have convergence in probability of the $\frac{Y_i^t(s)}{n}$ and underlying stochastic variables like $i_t(s) = \frac{I(t+s)}{n}$, $s_t(s) = \frac{S(t+s)}{n}$ and $r_t(s) = \frac{R(t+s)}{n}$ to deterministic variables $i_d(t+s), s_d(t+s), r_d(t+s)$ and $y_i^t(s)$, which satisfy the fractional Kermack-McKendrick equations by way of an integral equation, by almost sure convergence of martingale expressions $n^{-1} M_i^t(s)$:

$$\begin{pmatrix} s_t(s) \\ i_t(s) \end{pmatrix} = \begin{pmatrix} s_t(0) \\ i_t(0) \end{pmatrix} + n^{-1} \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} M_1^t(s) \\ M_2^t(s) \end{pmatrix} + \int_0^s \begin{pmatrix} -\lambda & 0 \\ \lambda & -\gamma \end{pmatrix} \begin{pmatrix} s_t(u) i_t(u) \\ i_t(u) \end{pmatrix} du \xrightarrow{P} \begin{pmatrix} a.s. also! \end{pmatrix}$$

$$\begin{pmatrix} s_d(t+s) \\ i_d(t+s) \end{pmatrix} = \begin{pmatrix} s_d(t) \\ i_d(t) \end{pmatrix} + \int_0^s \begin{pmatrix} -\lambda & 0 \\ \lambda & -\gamma \end{pmatrix} \begin{pmatrix} s_d(t+u) i_d(t+u) \\ i_d(t+u) \end{pmatrix} du \text{ (subscript } .d \text{ for deterministic):}$$

Caused by: $\sup_{0 \leq u \leq s} |n^{-1} M_i^t(u)| \xrightarrow{a.s.} 0$ and by that: $\sup_{0 \leq u \leq s} \left| \frac{Y_i^t(u)}{n} - y_i^t(u) \right| \xrightarrow{a.s.} 0, s \in [0, h]$

by a general theorem for centered Poisson processes.

Vitali's Convergence Theorem then not only ensures the convergence of expectations and variances of the 'basic' stochastic variables $i_t(s) = \frac{I(t+s)}{n}, s_t(s) = \frac{S(t+s)}{n}$ and $r_t(s) = \frac{R(t+s)}{n}$, but also of their linear combinations, as well as of their products and integral expressions.

b) Ch. 3: Accompanying Variation Processes $V(t)$ to zero mean martingales $M(t): Var(M(t)) = E(V(t))$

Translated to zero mean martingales: $M_i^t(s) = Y_i^t(s) - \int_0^s \lambda_i(t+u) ds, s \in [0, h]$, it can be proved:

For $[M_i^t(s)] = Y_i^t(s)$ and $\langle M_i^t(s) \rangle = \int_0^s \lambda_i(t+u) du$ and so for $M_i^t(s) = [M_i^t(s)] - \langle M_i^t(s) \rangle$ holds:

$$E([M_i^t(s)]) = E(\langle M_i^t(s) \rangle) = Var(M_i^t(s)) (= Y_i^t(s))$$