

Master's Thesis in requirement for partial completion

The University of Melbourne

Jacob Cumming

September 2023

Contents

1	Epidemiological Modelling	1
1.1	Deterministic ODE models	2
1.2	Stochastic models	4
1.3	Doob-Gillespie Algorithm	7
1.4	τ -leaping	7
2	Malaria and Malaria Models	9
2.1	Malaria	9
2.2	Malaria Models	10
3	Approximate Bayesian Computing	17
3.1	Summary Statistics	17
3.2	Discrepancy	17
3.3	Estimating the likelihood from the summary statistic	19
4	Gaussian Processes	21
4.1	Motivation and Preliminaries	21
4.2	Kernel Families	23
4.3	Differing mean functions	24
5	Stuff that could fix stuff	25
	REFERENCES	26

List of Figures

1.1	Some simple model schematics, with varying numbers of compartments: S (susceptable), E (exposed), I (infectious) and R (recovered). The force of infection λ_t is usually a function of I_t , depicted by the dashed red lines. μ and ν are natural birth and death rates respectively. γ is the rate of progression out of the infectious state. In each of these models the physical interpretation differs slightly. In the SIS and $SEIR$ models, it is the rate at which individuals move from infectious to susceptible again or into lifelong immunity, whereas in the SI with demography model, it can be interpreted as the increase to the rate of death attributable to disease induced mortality. σ is the rate of progression from a state of latent infection to becoming infectious.	2
1.2	Solutions to the ordinary differential equations describing the models depicted in Figure 1.1. The initial infectious population was $I_0 = 10$, with $S_0 = 990$. In the $SEIR$ model, $E_0 = R_0 = 0$. For all models $\beta = 0.4$. For the SIS and SI model with demography $\gamma = 1/4$. For the SI model with demography $\mu = 0.012$, and $\nu = 0.0012$. For the $SEIR$ model, $\gamma = 1/90$, and $\sigma = 1/2$	3
1.3	Exact stochastic simulations of the 3 different models using Algorithm 1. The parameters used were identical to those in Figure 1.2	7
2.1	The malaria lifecycle.	9
2.2	A simple Ross-Macdonald malaria model schematic, as described by Aron and May 1982. S_H and I_H are the number of susceptible and infected humans respectively, and S_M and I_M are the number of susceptible and infected mosquitos. The rate of human infection (λ_H) is dependant on I_M , and the rate of human infection (λ_M) is dependant on I_H	11
2.3	Diagram for $P. vivax$ model in a tropical setting described by White et al. 2016. S and I are the number of susceptible and infected humans and mosquitos (denoted by subscript M). $\lambda_H = mabI_M$ and $\lambda_M = ac(I_0 + I_L)$	12
2.4	Diagram for $P. vivax$ model described by Champagne et al. 2022. $I_{total} = I_0 + I_L$. Since the mosquito dynamics have been removed, λ now not has no dependencies on the number of infectious mosquitos.	14
3.1	Using a basic SIS model as in previous figure using ABC to generate from posterior beta, gamma	18

Chapter 1

Epidemiological Modelling

In order to study the behaviour and characteristics of disease spread and eradication, compartmental epidemiological models have been developed. They seek to simplify the dynamics of a disease down to a mathematically representable form. Inference on these models allow for an understanding of how the modelled disease spreads, and allows an assessment of how effective differing disease interventions (such as treatments or vaccinations) may be without the need for large long term trials. Models can also simulate various scenarios such as increases or decreases in viral transmission.

Simple compartmental disease models assume individuals can be only be in one of a finite number of states (which are called compartments). These compartments usually correspond to a state of disease. Some simple common compartments include:

- S - Susceptable: at risk of contracting the disease
- E - Exposed: contracted the disease but not yet transmitting it
- I - Infectious (also called Infected): at risk of transmitting the disease
- R - Recovered: neither at risk of contracting or transmitting the disease.

The number of people in each compartment at time t is a (possibly non-deterministic) function of time t , which we indicate as a subscript t (eg. S_t is the number of susceptibles at time t). Models are routinely described by the compartments they contain. For example, an SIS model, is a model with the susceptible and infectious compartments. Recovering from the infection leaves you susceptible to reinfection (for example most sexually transmitted diseases (Keeling and Rohani 2008, p. 56)), and is graphically depicted in Figure 1.1a.

Furthermore we can also include demography into a compartmental model. Diseases that infect the individual until the time of death such as bovine spongiform encephalopathy (BSE commonly known as mad cow disease) may be modelled using an SI model with demography (birth and death rates), as depicted in Figure 1.1b (Hagenaars, Donnelly, and Ferguson 2006).

Childhood diseases such as varicella (chickenpox) which give lifetime immunity after infection can be modelled using an $SEIR$ model (see Figure 1.1c), particularly when modelling a local outbreak setting (for example Zha et al. 2020 used the $SEIR$ model to model a school outbreak of varicella). Not including demography is usually appropriate when disease induced mortality is low.

The number (and names) of compartments can be extended and configured as needed, and compartments could be added for vaccinated individuals, quarantined individuals and so on.

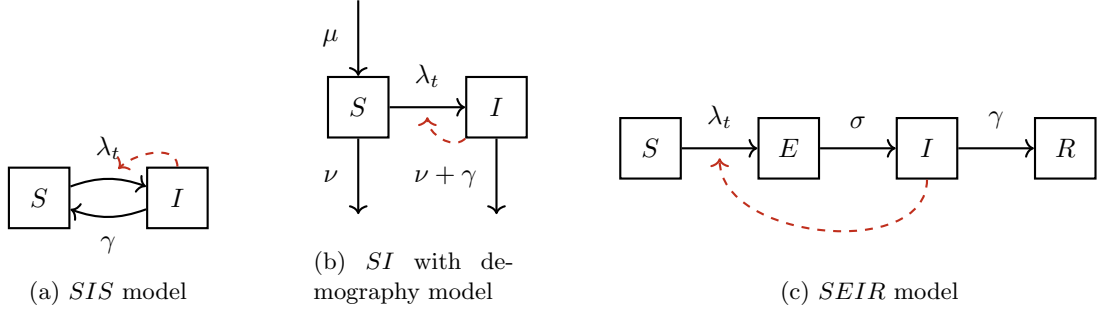


Figure 1.1: Some simple model schematics, with varying numbers of compartments: *S* (susceptible), *E* (exposed), *I* (infectious) and *R* (recovered). The force of infection λ_t is usually a function of I_t , depicted by the dashed red lines. μ and ν are natural birth and death rates respectively. γ is the rate of progression out of the infectious state. In each of these models the physical interpretation differs slightly. In the *SIS* and *SEIR* models, it is the rate at which individuals move from infectious to susceptible again or into lifelong immunity, whereas in the *SI* with demography model, it can be interpreted as the increase to the rate of death attributable to disease induced mortality. σ is the rate of progression from a state of latent infection to becoming infectious.

By convention N_t (often simply N in models with a closed population) is the total number of individuals in the model, the sum of all compartments.

1.1 Deterministic ODE models

Diseases are often simulated as deterministic ordinary differential equations. For the examples below we assume that the force of infection λ_t is proportional to the number of people in *I*, such that $\lambda_t := \beta \frac{I_t}{N_t}$. β can be interpreted as the average number of people that an individual interacts with per day in a way such that disease would be spread in that interaction per unit of time t . This is sometimes called the effective contact rate. Therefore since $\frac{I_t}{N_t}$ is the probability that a randomly selected individual is infectious, $\beta \frac{I_t}{N_t}$ can be interpreted as the average number of people that a person interacts with each day who are infectious in a way that they would pass on the infection. In different diseases β varies dramatically, as some diseases need prolonged exposure or sexual contact to transmit, whereas some are very highly transmittable, and so will have very low β . Implicitly there is also a (sometimes poor) assumption of complete uniformly random mixing of people. Note that for this thesis we assume that β is frequency dependent (people interact with the same number of people regardless of population size) as opposed to density dependent (people interact with a number of people proportional to population size, in which case $\lambda := \beta I_t$).

The *SIS* model depicted in Figure 1.1a, the ordinary different equations (ODEs) governing the model could be

$$\frac{dS_t}{dt} = -\lambda S_t + \gamma I_t = -\beta \frac{I_t}{N} S_t + \gamma I \quad (1.1)$$

$$\frac{dI_t}{dt} = \lambda S_t - \gamma I_t = \beta \frac{I_t}{N} S_t - \gamma I. \quad (1.2)$$

With a stated assumption that population size is closed, equation 1.1 fully describes the model.

The system of ODEs that describe the *SI* with demography model is

$$\frac{dS_t}{dt} = \mu N_t - \lambda_t S_t - \nu I_t = \mu N_t - \beta \frac{I_t}{N_t} S_t - \nu I_t \quad (1.3)$$

$$\frac{dI_t}{dt} = \lambda_t S_t - (\gamma + \nu) I_t = \beta \frac{I_t}{N_t} S_t - (\gamma + \nu) I_t. \quad (1.4)$$

Here it is important to note that N_t is not necessarily constant.

Finally, the system of ODEs that describe the *SEIR* model is

$$\frac{dS_t}{dt} = -\lambda_t S_t - \nu I_t = -\beta \frac{I_t}{N} S_t + \gamma I_t \quad (1.5)$$

$$\frac{dE_t}{dt} = \lambda_t S_t - \omega E_t = \beta \frac{I_t}{N} S_t - \omega I_t \quad (1.6)$$

$$\frac{dI_t}{dt} = \omega E_t - \gamma I_t \quad (1.7)$$

$$\frac{dR_t}{dt} = \gamma I_t \quad (1.8)$$

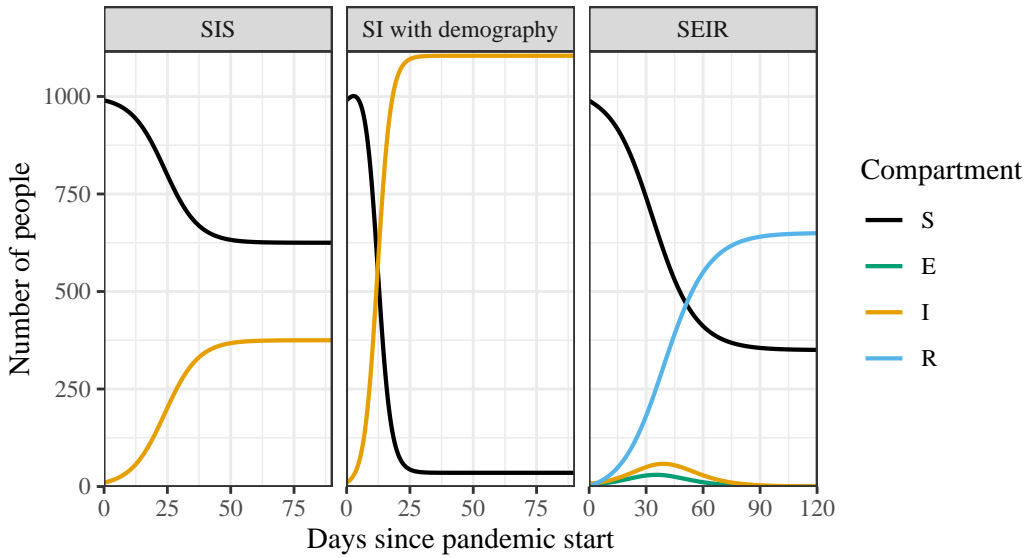


Figure 1.2: Solutions to the ordinary differential equations describing the models depicted in Figure 1.1. The initial infectious population was $I_0 = 10$, with $S_0 = 990$. In the *SEIR* model, $E_0 = R_0 = 0$. For all models $\beta = 0.4$. For the *SIS* and *SI* model with demography $\gamma = 1/4$. For the *SI* model with demography $\mu = 0.012$, and $\nu = 0.0012$. For the *SEIR* model, $\gamma = 1/90$, and $\sigma = 1/2$.

After specifying the initial for each compartment the ordinary differential equations have a deterministic output, such as in 1.2.

1.2 Stochastic models

Motivating the form of the stochastic model

To establish a relationship between the deterministic and stochastic disease models, we first need to establish Poisson point processes and their properties.

Definition 1.1 (Poisson Point Process). $\{\mathcal{N}(t)\}_{t \geq 0}$ is a (stationary) Poisson point process with intensity λ if

1. $\mathcal{N}(0) = 0$
2. $\mathcal{N}(t_1) - \mathcal{N}(t_0), \mathcal{N}(t_2) - \mathcal{N}(t_1), \dots, \mathcal{N}(t_n) - \mathcal{N}(t_{n-1})$ are independent for $0 \leq t_0 < t_1 < \dots < t_{n-1} < t_n$
3. $\mathcal{N}(t_2) - \mathcal{N}(t_1) \sim \text{Pois}(\lambda(t_2 - t_1)), 0 \leq t_1 < t_2$.

Deterministic ODE models are appropriate to study a kind of aggregate disease spread behaviour, and well approximate real world behaviour when the numbers in each compartment are large, however at the start of an epidemic, when the number of infected individuals is small the behaviour of the epidemic may vary significantly. It is possible that if the average number of people that an infectious individual infects near the beginning of the epidemic (formally referred to as R_0) is close to 1, then the disease may die out or become stable. Under the deterministic SIS model described by equations 1.1 and 1.2, consider the model at time t^* the instantaneous rate at which S is decreasing is $\beta \frac{I_{t^*}}{N} S_{t^*}$. In other words, one individual leaves the S compartment every $\beta \frac{I_{t^*}}{N} S_{t^*}$ units of time. We can consider a Poisson point process $\{\mathcal{N}_1(t - t^*)\}_{t \geq t^*}$ with intensity $\beta \frac{I_{t^*}}{N} S_{t^*}$ corresponding to the count of the number of individuals who have left S and entered I t units of time since t^* .

$$\frac{d\mathbb{E}(\mathcal{N}_1(t^*))}{dt^*} = \lim_{\delta \rightarrow 0} \frac{\mathbb{E}(\mathcal{N}(t^* + \delta) - \mathcal{N}(t^*))}{\delta} = \frac{\beta \frac{I_{t^*}}{N} S_{t^*}(t^* + \delta - t^*) \beta \frac{I_{t^*}}{N} S_{t^*}}{\delta} = \beta \frac{I_{t^*}}{N} S_{t^*}.$$

Under the same deterministic formulation of the model, the instantaneous rate into S at t^* is γI_{t^*} . Therefore as above we can construct a Poisson point process $\{\mathcal{N}_2(t - t^*)\}_{t \geq t^*}$ with rate γI_{t^*} describing the number of recoveries from I to S , with $\frac{d\mathbb{E}(\mathcal{N}_2(t^*))}{dt^*} = \gamma I_{t^*}$. Combining the two processes, we can see that the rate of change in the average number of people in S is

$$\frac{d\mathbb{E}(\mathcal{N}_2(t^*) - \mathcal{N}_1(t^*))}{dt^*} = \frac{d\mathbb{E}(\mathcal{N}_2(t^*)) - d\mathbb{E}(\mathcal{N}_1(t^*))}{dt^*} = -\beta \frac{I_t}{N} S_t + \gamma I = \frac{dS_t}{dt}.$$

Therefore we can create a stochastic model where the local average in each compartment matches the behaviour of the ODE model at the same state. We do this first by formulating the model as a random vector $\{\mathbf{C}_t\}_{t \geq 0} = \{C_1(t), C_2(t), \dots, C_n(t)\}_{t \geq 0}$ where $C_i : \mathbb{R} \rightarrow \mathbb{N} \cup \{0\}$, is the number of people in compartment C_i , and for any fixed t , $\{C_1(t), C_2(t), \dots, C_n(t)\}$ is a random variable describing the state of the model. For example in a model with S and I compartments, $\{\mathbf{C}_t\}_{t \geq 0} := \{S_t, I_t\}_{t \geq 0}$. $\{\mathbf{C}_t\}_{t \geq 0}$ is a continuous time Markov chain (see Definition ??) with transition kernel corresponding to the rates of the model. For example, in the SI model with demography in Figure 1.1b the transition rates are:

- $\{s, i\}$ to $\{s + 1, i\}$ has rate $\mu(s + i)$
- $\{s, i\}$ to $\{s - 1, i\}$ has rate νs

- $\{s, i\}$ to $\{s - 1, i + 1\}$ has rate $\beta \frac{i}{i+s} s$
- $\{s, i\}$ to $\{s, i - 1\}$ has rate $(\nu + \gamma)i$.

We can interpret each transition as a separate events, each behaving as independent Poisson point processes until the time of the first transition. Therefore at time t^* we have the Poisson point processes:

- $\{\mathcal{E}_1(t)\}_{t \geq 0}$: the number of births into S after time t^* with intensity μN_{t^*}
- $\{\mathcal{E}_2(t)\}_{t \geq 0}$: the number of deaths in S after time t^* with intensity νS_{t^*}
- $\{\mathcal{E}_3(t)\}_{t \geq 0}$: the number of infections after time t^* with intensity $\beta \frac{I_{t^*}}{N_{t^*}} S_{t^*}$
- $\{\mathcal{E}_4(t)\}_{t \geq 0}$: the number of deaths from I after time t^* with intensity $(\nu + \gamma)I_{t^*}$.

Theorem 1.1 (Sums of Independent Poisson Point Processes). *Given independent Poisson point processes $\{\mathcal{N}_1(t)\}_{t \geq 0}, \{\mathcal{N}_2(t)\}_{t \geq 0}, \dots, \{\mathcal{N}_n(t)\}_{t \geq 0}$, with intensities $\lambda_1, \lambda_2, \dots, \lambda_n$,*

$$\{\mathcal{N}(t)\}_{t \geq 0} := \{\mathcal{N}_1(t) + \mathcal{N}_2(t) + \dots + \mathcal{N}_n(t)\}_{t \geq 0}$$

is a Poisson point process with intensity $\lambda_1 + \lambda_2 + \dots + \lambda_n$.

Proof. We show that $\{\mathcal{N}(t) := \{\mathcal{N}_1(t) + \mathcal{N}_2(t) + \dots + \mathcal{N}_n(t)\}_{t \geq 0}\}$ meets each component of Definition 1.1.

1. $\mathcal{N}(0) := \mathcal{N}_1(0) + \mathcal{N}_2(0) + \dots + \mathcal{N}_n(0) = 0$ since $\mathcal{N}_i(0) = 0$ by definition of a Poisson point process.
2. We show that $\mathcal{N}(t_1) - \mathcal{N}(t_0), \mathcal{N}(t_2) - \mathcal{N}(t_1), \dots, \mathcal{N}(t_n) - \mathcal{N}(t_{n-1})$ with $0 \leq t_0 < t_1 < \dots < t_n$ are independent.

$$\mathcal{N}(t_i) - \mathcal{N}(t_{i-1}) = \underbrace{[\mathcal{N}_1(t_i) - \mathcal{N}_1(t_{i-1})]}_{X_{i1}} + \underbrace{[\mathcal{N}_2(t_i) - \mathcal{N}_2(t_{i-1})]}_{X_{i2}} + \dots + \underbrace{[\mathcal{N}_n(t_i) - \mathcal{N}_n(t_{i-1})]}_{X_{in}}$$

X_{ik} is independent of $X_{j\ell}$ for $k \neq \ell$ since \mathcal{N}_k and \mathcal{N}_ℓ are independent processes. X_{ik} is independent of X_{jk} for $i \neq j$ by the second property of Definition 1.1. Therefore all X_{ik} are independent of $X_{j\ell}$ for $i \neq j$, and all j, k . Hence $\mathcal{N}(t_1) - \mathcal{N}(t_0), \mathcal{N}(t_2) - \mathcal{N}(t_1), \dots, \mathcal{N}(t_n) - \mathcal{N}(t_{n-1})$ with $0 \leq t_0 < t_1 < \dots < t_n$ are independent.

3. For fixed $t_1 < t_2$, and $i \in \{1, 2, \dots, n\}$,

$$\mathcal{N}_i(t_2) - \mathcal{N}_i(t_1) \sim \text{Pois}((t_2 - t_1)\lambda_i).$$

Consider the associated moment generating function of $\mathcal{N}_i(t_2) - \mathcal{N}_i(t_1)$,

$$M_i(z) := \exp(\lambda_i(t_2 - t_1)(\exp(z) - 1)).$$

Therefore the moment generating function of

$$\mathcal{N}(t_2) - \mathcal{N}(t_1) = [\mathcal{N}_1(t_2) - \mathcal{N}_1(t_1)] + [\mathcal{N}_2(t_2) - \mathcal{N}_2(t_1)] + \dots + [\mathcal{N}_n(t_2) - \mathcal{N}_n(t_1)]$$

is

$$M(z) := \prod_{i=1}^n M_i(z) = \exp[(\lambda_1(t_2 - t_1) + \lambda_2(t_2 - t_1) + \dots + \lambda_n(t_2 - t_1))(\exp(z) - 1)].$$

Therefore $\mathcal{N}_1(t) + \mathcal{N}_2(t) + \dots + \mathcal{N}_n(t) \sim \text{Pois}((\lambda_1 + \lambda_2 + \dots + \lambda_n)t)$ by the uniqueness of the moment generating function.

□

Theorem 1.2 (Time to First Event in Poisson Point Process). *Given a Poisson point process $\{\mathcal{N}(t)\}_{t \geq 0}$ with intensity λ , let $\tau = \inf\{t | \mathcal{N}(t_0 + t) - \mathcal{N}(t_0) = 1, t > 0\}$. $\tau \sim \text{Exp}(\lambda)$ for $t_0 \geq 0$*

Proof.

$$\Pr(\tau > x) = \Pr(\mathcal{N}(t_0 + x) - \mathcal{N}(t_0) = 0) = \frac{(\lambda x)^0 e^{-\lambda x}}{0!} = e^{-\lambda x}$$

□

By Theorem 1.1 and Theorem 1.2,

$$\{\mathcal{E}(t)\}_{t \geq 0} := \{\mathcal{E}_1(t) + \mathcal{E}_2(t) + \mathcal{E}_3(t) + \mathcal{E}_4(t)\}_{t \geq 0}$$

is a Poisson point process with intensity

$$\mu N_{t^*} + \nu S_{t^*} + \beta \frac{I_{t^*}}{N_{t^*}} S_{t^*} + (\nu + \gamma) I_{t^*},$$

and the time to the next event is random variable distributed

$$\text{Exp}(\mu N_{t^*} + \nu S_{t^*} + \beta \frac{I_{t^*}}{N_{t^*}} S_{t^*} + (\nu + \gamma) I_{t^*}).$$

Theorem 1.3 (Probability of i th Poisson Process Generating the Next Event). *Consider independent Poisson point processes*

$$\{\mathcal{N}_1(t)\}_{t \geq 0}, \{\mathcal{N}_2(t)\}_{t \geq 0}, \dots, \{\mathcal{N}_n(t)\}_{t \geq 0}$$

having intensities $\lambda_1, \lambda_2, \dots, \lambda_n$. For fixed t_0 , let $\tau_i := \inf\{t | \mathcal{N}(t_0 + t) - \mathcal{N}(t_0) = 1\}$. Then

$$\Pr(\min_i \tau_i = \tau_j) = \frac{\lambda_j}{\sum_{i=1}^n \lambda_i}.$$

Proof. By Theorem 1.2, $\tau_i \sim \text{Exp}(\lambda_i)$. Therefore

$$\begin{aligned} \Pr(\min_i \tau_i = \tau_j) &= \int_0^\infty \Pr(\{\tau_i = x\} \cup \bigcup_{j \neq i} \{\tau_j > x\}) dx \\ &= \int_0^\infty \Pr(\{\tau_i = x\} \cup \bigcup_{j \neq i} \{\tau_j > x\}) dx \\ &= \int_0^\infty \Pr(\tau_i = x) \times \prod_{j \neq i} \Pr(\tau_j > x) dx && \text{(by independence)} \\ &= \int_0^\infty \lambda_i \exp(-\lambda_i x) \times \prod_{j \neq i} \exp(-\lambda_j x) dx \\ &= \lambda_i \int_0^\infty \exp(-(\sum_{j=1}^n \lambda_j) x) dx \\ &= \lambda_i \left[\frac{\exp(-(\sum_{j=1}^n \lambda_j) x)}{\sum_{j=1}^n \lambda_j} \right]_0^\infty \\ &= \frac{\lambda_i}{\sum_{j=1}^n \lambda_j} \end{aligned}$$

□

1.3 Doob-Gillespie Algorithm

All of this leads naturally to a common method of simulating the stochastic model. The Doob-Gillespie algorithm (often just called the Gillespie algorithm) is an algorithm that simulates a stochastic realisation of a model given a set of starting conditions.

Algorithm 1 The Doob-Gillespie Algorithm

- 1: Initialise time $t \leftarrow 0$ and initial state of the model $\mathbf{C}(0) := \{C_1(0), C_2(0), \dots, C_n(0)\}$
 - 2: **while** termination condition not met **do**
 - 3: Calculate intensities λ_i for all possible events \mathcal{E}_i
 - 4: Calculate total intensity $\lambda = \sum_i \lambda_i$
 - 5: Generate $\Delta t \sim \text{Exp}(\lambda)$
 - 6: Choose event E_i with probability $\frac{\lambda_i}{\lambda}$
 - 7: Update time $t \leftarrow t + \Delta t$
 - 8: Update state of $\mathbf{C}(t + \delta t) \leftarrow \mathbf{C}(t) + \text{change in state due to event } \mathcal{E}_i$
 - 9: **end while**
-

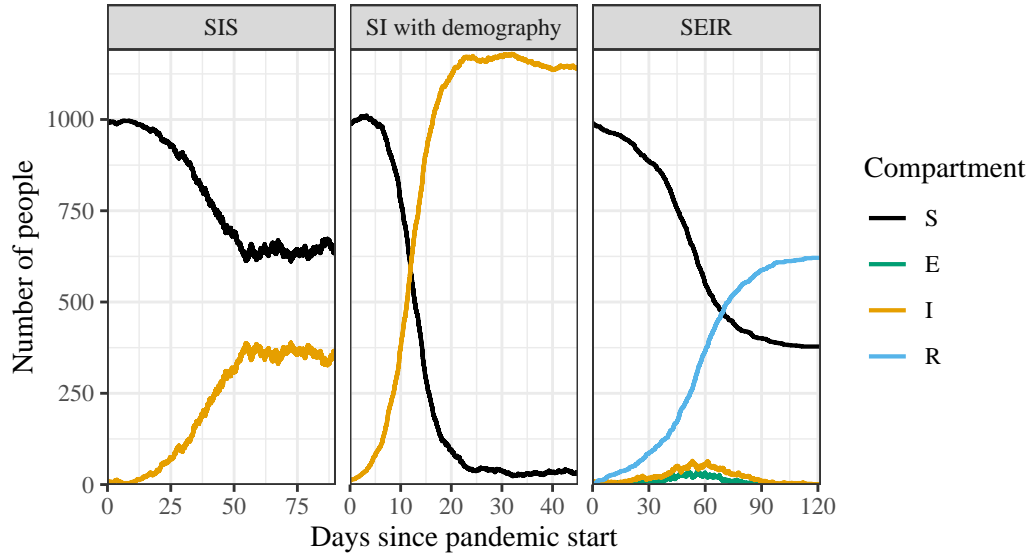


Figure 1.3: Exact stochastic simulations of the 3 different models using Algorithm 1. The parameters used were identical to those in Figure 1.2

1.4 τ -leaping

τ -leaping exploits the local Poisson point process like behaviour of epidemiological models. Consider the SIS model, when $S_t = I_t = 10000$. Events happen at a very high rate, meaning the Δt found in each step of the Doob-Gillespie algorithm will be very small, but the rates also change a negligible amount after each event (compare $\gamma \times 10000$ to $\gamma \times 10001$ or $\gamma \times 9999$). Therefore we can approximate the number of events in a short time period τ as a Poisson point process

with the total intensity $\lambda = \sum_i \lambda_i$ at time t , with the probability of any one event having the same probability as above of $\frac{\lambda_i}{\lambda}$. Therefore we have the following algorithm.

Algorithm 2 τ -Leaping Algorithm

- 1: Initialise time $t \leftarrow 0$ and initial state of the model $\mathbf{C}(0) := \{C_1(0), C_2(0), \dots, C_n(0)\}$
 - 2: **while** termination condition not met **do**
 - 3: Calculate intensities λ_i for all possible events \mathcal{E}_i
 - 4: Calculate total intensity $\lambda = \sum_i \lambda_i$
 - 5: Choose a suitable time step τ (this can be deterministic or adaptive)
 - 6: Calculate Poisson random variable $X \sim \text{Poisson}(\lambda)$
 - 7: **for** i in 1 to X **do**
 - 8: Choose event \mathcal{E}_i with probability $\frac{\lambda_i}{\lambda}$
 - 9: Update state of $\mathbf{C}(t + \tau) \leftarrow \mathbf{C}(t) +$ change in state due to event \mathcal{E}_i
 - 10: **end for**
 - 11: Update time $t \leftarrow t + \tau$
 - 12: **end while**
-

Chapter 2

Malaria and Malaria Models

2.1 Malaria

Malaria kills around 600,000 people each year, with over 75% of deaths occurring in children under 5 years old (World Health Organization 2022).

Plasmodium Life Cycle

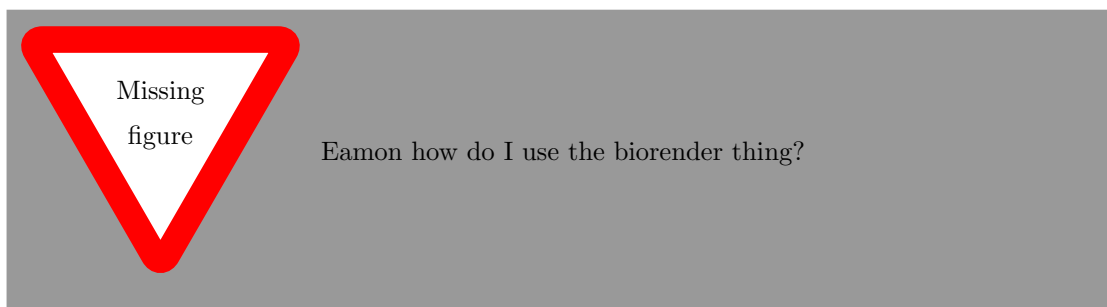


Figure 2.1: The malaria lifecycle.

Malaria is a vector borne disease, needing both human (or other vertebrate) and mosquito hosts to complete its lifecycle (Figure 2.1). Six species of the unicellular parasite are able to infect humans (Milner 2018). Although *Plasmodium falciparum* is responsible for around 90% of total human malaria deaths, outside of Africa *Plasmodium vivax* is the leading cause of malaria infection (Zekar and Sharman 2023; Adams and Mueller 2017). Sporozoites (a stage of the malaria parasite) enter the human blood stream via the skin after the female mosquito has a blood meal. From the blood stream they proceed to enter into the liver. Once a hepatocyte (liver cell) is invaded, the parasite will undergo asexual reproduction into up to 40,000 merozoites per hepatocyte, which are released into the blood stream. These merozoites then bind to, and invade erythrocytes (red blood cells), once again reproducing 16-32 fold in a process called schizogony. At this point, the erythrocyte membrane is ruptured, allowing for *Plasmodium* to invade new erythrocytes. Eventually, the merozoites undergo sexual differentiation, resulting in the sequestration and maturation of male and female gametocytes in the bone marrow, until they are released into the blood stream to be consumed by a mosquito during a blood feed where

it matures into sporozoites ready to reinfect a new vertebrate host when the mosquito next takes a blood feed (Cowman et al. 2016).

Illness, Treatment, and Immunity

The most common symptom of malaria infection in persons without natural or acquired immunity is fever. After treatment, fever will usually subside over a few days. In severe cases, malaria can lead to anemia, cerebral malaria (coma), and respiratory distress (Cowman et al. 2016).

In a population with stable malarial infection, immunity increases with age, with the proportion of severe cases negligible after age 10, and asymptomatic infection being the dominant infection type beyond age 15 (Cowman et al. 2016).

Control and Eradication

Widespread use of DDTs in the mid 20th century led to significant successes in some countries towards the control and eradication of malaria. In the 1980s and 1990s, drug resistant malaria led to a doubling of malaria-attributable death. Currently, the control techniques include insecticide treated bed nets, and a mixture of antimalarial drugs (Cowman et al. 2016).

Plasmodium vivax

Unlike *P. falciparum*, *P. vivax* has hypnozoites, which are a dormant liver stage of the parasite. These can remain dormant for weeks and even months, leading to recurrent infections and illness, possibly until the conditions for transmission are more favourable. In subtropical/temperate areas, the incubation periods can be between 8-12 months, compared to 3-4 weeks in tropical regions. Price et al. 2020. *P. vivax* also has lower levels of the blood stage parasite during infection, which means diagnosis is more difficult, and it has an increased proportion of asymptomatic cases (Adams and Mueller 2017).

It is likely that death and severe disease attributable to *P. vivax* has been traditionally underestimated. In view of recent evidence, the old notion that *P. vivax* is benign has become untenable (Cowman et al. 2016).

Motivating Malaria Models

Levels of asymptomatic cases and latent parasite (in the case of *P. vivax*) are impossible or difficult to experimentally determine without mass testing. By creating a model of the disease, and calibrating the model so that it simulates symptomatic case levels reported by health authorities, it is possible to estimate these previously ‘hidden’ levels. Furthermore, modelling malaria allows for modelling the effect of public health interventions such as mass treatment or testing, in order to determine an estimate of how effective the intervention may be, before large amounts of money are spent on trials.

2.2 Malaria Models

Ross-Macdonald

Modelling malaria presents an additional challenge, as the disease is transmitted from mosquito to human and human to mosquito, rather than having direct human to human transmission.

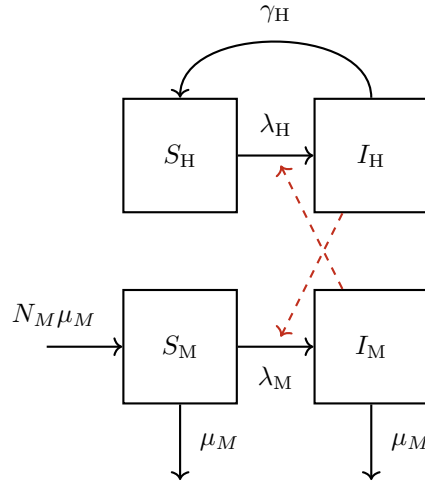


Figure 2.2: A simple Ross-Macdonald malaria model schematic, as described by Aron and May 1982. S_H and I_H are the number of susceptible and infected humans respectively, and S_M and I_M are the number of susceptible and infected mosquitos. The rate of human infection (λ_H) is dependant on I_M , and the rate of human infection (λ_M) is dependant on I_H .

The most simple Ross-Macdonald model is depicted in figure 2.2. The ODEs for this model are

$$\begin{aligned}\frac{dS_H}{dt} &= \gamma_H I_H - b T_{HM} I_M \frac{S_H}{N_H} \\ \frac{dI_H}{dt} &= b T_{HM} I_M \frac{S_H}{N_H} - \gamma_H I_H \\ \frac{dS_M}{dt} &= N_M \mu_M + \gamma_M I_M - b T_{MH} S_M \frac{I_H}{N_H} - S_M \mu_M \\ \frac{dI_M}{dt} &= b T_{MH} S_M \frac{I_H}{N_H} - \gamma_M I_M\end{aligned}$$

where b is the biting rate per mosquito, and T_{HM} is the probability of tranmission to a human given a bite by an infectious mosquito, with T_{MH} being vice-versa. Note that it is $\frac{I_H}{N_H}$ in the mosquito dynamics. Biologically this is assuming the number of blood meals a mosquito takes per day is invariant to the size of the human population. Mosquitos don't 'recover' from malaria due to their short lifespans, but the births and deaths are mathematically equivalent to assuming that the rate of 'recovery' amongst mosquitos is $\mu_M I_M$ per unit time, with no population dynamics.

A Ross-Macdonald style model simplifies the lifecycle of malaria to the following four steps (Smith et al. 2012):

1. Malaria is transmitted to human (or vertebrate) via a blood feed.
2. Malaria proliferates in the human host until it circulates in the peripheral blood
3. A mosquito then takes a blood feed, ingesting the pathogen
4. Malaria develops within the mosquito host, progressing to its salivary glands, able to infect a human.

Models of *P. Vivax* Malaria

White Model

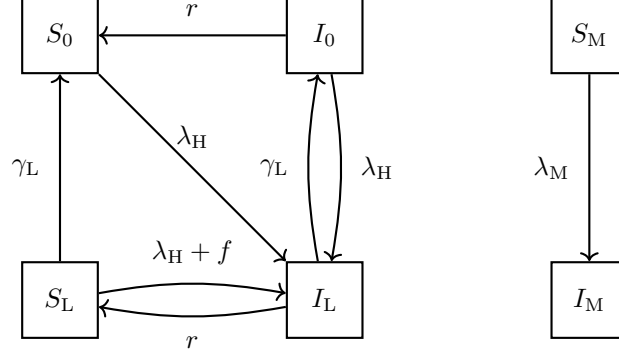


Figure 2.3: Diagram for *P. vivax* model in a tropical setting described by White et al. 2016. S and I are the number of susceptible and infected humans and mosquitos (denoted by subscript M). $\lambda_H = mabI_M$ and $\lambda_M = ac(I_0 + I_L)$

The White model - described in White et al. 2016 (tropical model) and depicted in Figure 2.3 - is characterised by the following ordinary differential equations:

$$\begin{aligned}
 \frac{dS_0}{dt} &= -\lambda_H S_0 + r I_0 + \gamma_L S_L \\
 \frac{dI_0}{dt} &= -\lambda_H I_0 - r I_0 + \gamma_L I_L \\
 \frac{dS_L}{dt} &= -\lambda_H S_L + r I_L - f S_L - \gamma_L S_L \\
 \frac{dI_L}{dt} &= \lambda_H (S_0 + I_0 + S_L) - r I_L + f S_L - \gamma_L I_L \\
 \frac{dS_M}{dt} &= g - \lambda_M (p S_M - (1-p) I_M) - g S_M \\
 \frac{dI_M}{dt} &= \lambda_M (p S_M - (1-p) I_M) - g I_M. \quad (I_0 + I_L = \text{total number of bloodstage infections})
 \end{aligned}$$

It modifies the Ross-Macdonald models, to capture the differences in disease progression between *P. vivax* and *P. falciparum*. In particular, the White model includes the dormant liver stage that is unique to *P. vivax*.

The model is comprised of six compartments:

1. **S_0 (Susceptible Individuals - No Latent Hypnozoite Liver Stage Infection):** People in this compartment have no form of malarial infection. These people are susceptible to new malarial infections, and are infected into compartment I_L (with both blood and liver stage parasites) with rate λ_H .
2. **I_L (Infected Individuals - Both Blood Stage and Latent Hypnozoite Liver Stage Infection):** Individuals in this compartment have both an active blood-stage infection,

and latent hypnozoite infection in the liver. They can progress to either I_0 through the clearance of liver stage infection with rate γ_L , or to S_L through clearance of blood stage infection with rate r .

3. I_0 (**Infected Individuals - Blood-Stage Infection Only**): Those in this compartment have a blood-stage infection with no latent hypnozoite infection in the liver. They are reinfected into I_L with rate λ_H , relapse with rate f . Blood-stage infection is cleared (moving into compartment S_0) with rate r .
4. S_L (**Susceptable Individuals - Blood-Stage Infection Only**): Those in this compartment have latent hypnozoite infection in the liver without blood-stage infection. They get novel infection through a mosquito bite into I_L with rate λ_H , or hypnozoite activation with rate f . This means that those in S_L move to compartment I_L with total rate $\lambda_H + f$. Alternatively the hypnozoites are cleared from the liver (moving to compartment S_0) with rate γ_L .
5. S_M (**Susceptable Mosquitoes**): Susceptable mosquitoes become infectious at rate $\lambda_M p$. They die at rate $g + \lambda_M(1 - p)$. Since there is a constant mosquito population assumption, mosquitoes are born into this state at rate $g + \lambda_M$.
6. I_M (**Infectious Mosquitoes**): Infectious mosquitos die at rate $g + \lambda_M(1 - p)$.

$\lambda_H := mabI_M$ where m is the number of mosquitos per human (held constant since there is no birth or death in the human dynamics), a is the mosquito biting rate, and b is the probability that a human bitten by an infectious mosquito develops an infection.

$\lambda_M := ac(I_0 + I_L)$ where a is defined above, and c is the probability that a mosquito bite on an infectious mosquito causes the mosquito to become infectious. g can be interpreted as the natural birth/death rate for mosquitos. p is then the proportion of mosquitos that survive long enough after the initial infection that the parasite matures enough in the mosquito before becoming infectious to new susceptible humans. Under the assumption that time until parasite transmissability after infection in a mosquito is a constant n days, and that mosquitoes naturally die at rate g , $p = e^{-gn}$. To see this let $V \sim \text{Exp}(g)$, represent the lifespan of the mosquito. $\Pr(V > n) = 1 - F_V(n) = 1 - (1 - e^{-gn}) = e^{-gn}$.

$\lambda_M(1 - p)$ can be interpreted as an additional rate of death, where of the mosquitos that would develop malaria after a bite, a proportion $1 - p$ die instantly. This applies to both the susceptible and infectious mosquitoes. Presumably this approximates a model where mosquitoes are moved to an 'exposed' compartment for n time, after initial infection, however no justification is given in (White et al. 2016) for this additional parameter n . A more straightforward SI model could be constructed that absorbs c and n into the single parameter c^* , such that it becomes the proportion of mosquito bites on blood stage infectious humans that result in mosquito infection where the mosquito does not die before becoming infectious. With steady mosquito population, the mosquito dynamics would now be characterised by

$$\frac{dI_M}{dt} = \lambda_M^* S_M - gI_M \quad \text{where } \lambda_M^* := ac^*(I_0 + I_L).$$

By modelling both liver and bloodstage infection, blood stage infections from relapses can be captured in the dynamics, meaning it is possible to analyse case number data that may be confounded by relapses as well as novel infections.

This model does not account for continual depletion of liver stage parasites which would vary the rate of relapse over time (through clearance or relapse). It also does not directly model any interventions or case importations. The lack of population dynamics means the model may only

be useful on a small time scale. Finally, it doesn't account for any importation of disease from an outside area, so if $S_0 = 1$, *P. vivax* is presumed permanently eradicated.

Champagne Model

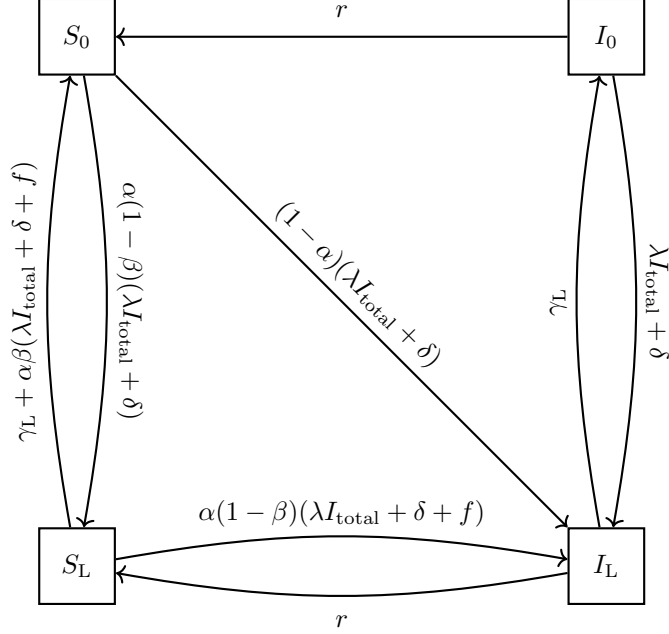


Figure 2.4: Diagram for *P. vivax* model described by Champagne et al. 2022. $I_{\text{total}} = I_0 + I_L$. Since the mosquito dynamics have been removed, λ now not has no dependencies on the number of infectious mosquitoes.

The Champagne model - described in (Champagne et al. 2022) and diagrammatically depicted in figure 2.4 - both simplifies and extends the White model. The model assumes human to human transmission, removing mosquito dynamics, and extends it by adding in a rate of imported cases and treatment of malarial infection. It is characterised by the system of ordinary differential equations

$$\begin{aligned} \frac{dI_L}{dt} &= (1 - \alpha)(\lambda I_{\text{total}} + \delta)(S_0 + S_L) + (\lambda I_{\text{total}} + \delta)I_0 + (1 - \alpha)fS_L - \gamma_L I_L - rI_L \\ \frac{dI_0}{dt} &= -(\lambda I_{\text{total}} + \delta)I_0 + \gamma_L I_L - rI_0 \\ \frac{dS_L}{dt} &= -(1 - \alpha(1 - \beta))(\lambda I_{\text{total}} + \delta + f)S_L + \alpha(1 - \beta)(\lambda I_{\text{total}} + \delta)S_0 - \gamma_L S_L + rI_L \\ \frac{dS_0}{dt} &= -(1 - \alpha\beta)(\lambda I_{\text{total}} + \delta)S_0 + (\lambda I_{\text{total}} + \delta)\alpha\beta S_L + \alpha\beta fS_L + \gamma_L S_L + rI_0 \end{aligned}$$

where $I_{\text{total}} := I_0 + I_L$.

The compartments have the same interpretation as in the White model, however the rates between compartments are significantly modified.

The new parameters are λ : the rate of infection, δ : importation rate, α : proportion of those infected but cleared of blood stage infections (through treatment), and β : a further proportion are also cleared of liver stage parasites, given that they were also cleared of blood stage infection (radical cure). In other words, the proportion of infected individuals $\alpha\beta$ are completely cured from liver and blood stage parasites. The model assumes treatment clears infection instantaneously. Individuals in S_L who relapse or get a new infection are assumed to be cured with the same proportions as new infections from S_0 , but individuals in I_0 who are superinfected are assumed not to seek treatment.

In contrast to the White model, the Champagne model allows analysis of potential treatment interventions, or how much of an impact limiting the importation rate might have on case numbers (through border control/testing). Although the lack of mosquito dynamics simplifies the model and it's running, it is unrealistic. The model still has some of the same problems as the White model, such as not incorporating hypnozoite depletion rates and a lack of population dynamics, meaning all analytic results are done assuming the system is at equilibrium.

Chapter 3

Approximate Bayesian Computing

Let $(\theta_i)_{i \in \{1, \dots, n\}}$ be a set of parameters with prior distribution π , and $(x_j)_{j \in \{1, \dots, m\}}$ be the set of realised data, generated using $(\theta_i)_{i \in \{1, \dots, n\}}$. The most basic approximate Bayesian computation algorithm is such:

1. **Generate** $(\hat{\theta}_i)_{i \in \{1, \dots, n\}} \sim \pi$
2. **Using** $(\theta_i)_{i \in \{1, \dots, n\}}$, **regenerate the data** $(\hat{x}_j)_{j \in \{1, \dots, m\}}$ **using the parameter set** $(\hat{\theta}_i)_{i \in \{1, \dots, n\}}$
3. **If** $(\hat{x}_j)_{j \in \{1, \dots, m\}} = (x_j)_{j \in \{1, \dots, m\}}$, **accept** $(\hat{\theta}_i)_{i \in \{1, \dots, n\}}$, **else regenerate** $(\hat{\theta}_i)_{i \in \{1, \dots, n\}}$ **and go to step 2**

For non-discrete cases

Generate parameter(s) from your prior.

Generate the data.

If it matches (or is close enough), add that parameter to your set of accepted parameters.

This set will approximate the posterior distribution.

Drawbacks of this approach. Very difficult to estimate multiple parameters to the one outcome parameters. Or if you have a single parameter, generating multiple observations, unlikely to match all of them. Same with multiple parameters to multiple outputs

3.1 Summary Statistics

Distribution of summary statistic

Should be relatively normally distributed

3.2 Discrepancy

Distribution of discrepancy

—Normal - normal— or something so should be normalish?

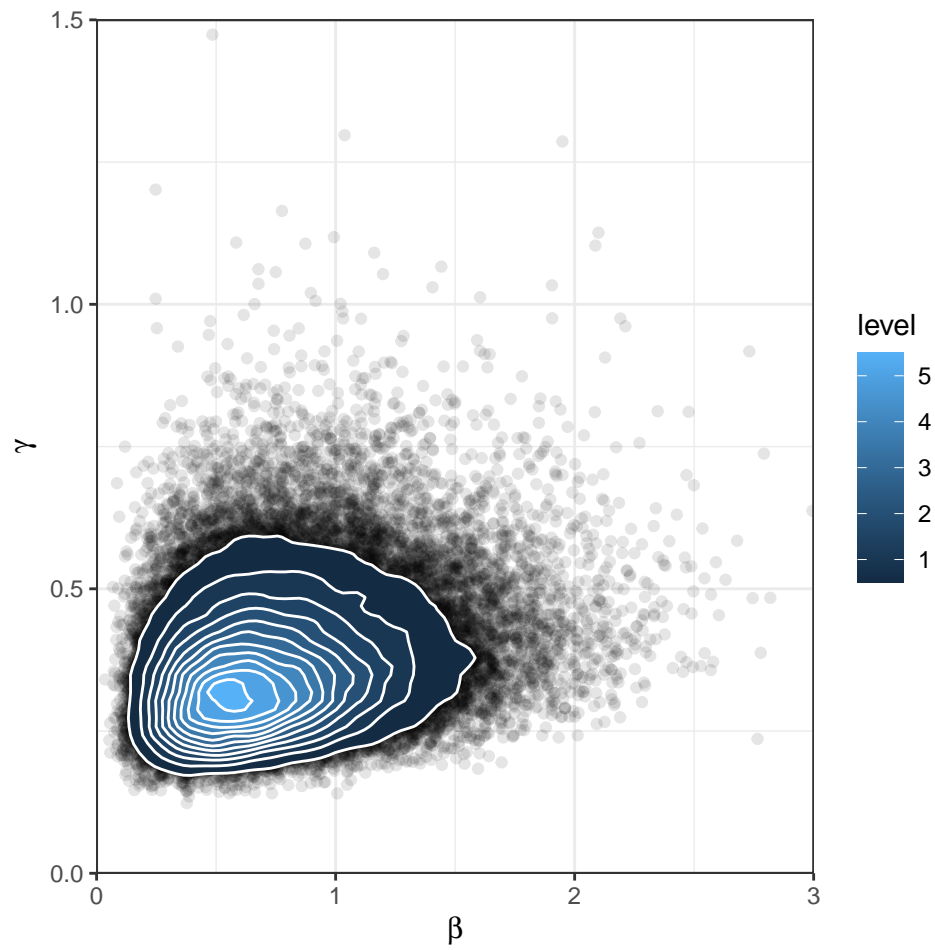


Figure 3.1: Using a basic SIS model as in previous figure using ABC to generate from posterior beta, gamma

3.3 Estimating the likelihood from the summary statistic

Chapter 4

Gaussian Processes

- Prove all pos-sem is cov
- expand on intro to motive the section
- prove exponential quad kern is pos semi-def

4.1 Motivation and Preliminaries

Our consideration of Gaussian processes is motivated by estimating functions that cannot be expressed in a closed form the context of estimating unknown functions from limited samples of that function. This is particularly helpful in disease modelling, when we use a Gaussian process prior to estimate the discrepancy of the summary statistics

Definition 4.1 (Kernel Function (symmetric, stationary)). *Any function $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is a kernel function. If $k(x, x') = k(x', x)$, the kernel is called symmetric.*

Definition 4.2 (Gaussian Process). *A collection of random variables $\{f(x)\}_{x \in \mathcal{X}}$ (where x may be a vector) is a Gaussian process if all finite dimensional distributions are multivariate normal distributed. That is, there is a function $m : \mathcal{X} \rightarrow \mathbb{R}$ and symmetric kernel $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ such that for all finite sets $\mathbf{x} := \{x_1, x_2, \dots, x_n\} \subset \mathcal{J}$, with $f(\mathbf{x}) := [f(x_1), f(x_2), \dots, f(x_n)]^T$*

$$f(\mathbf{x}) \sim \text{MVN} \left(\begin{bmatrix} m(x_1) \\ m(x_2) \\ \vdots \\ m(x_n) \end{bmatrix}, \mathbf{K} = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ k(x_n, x_1) & \dots & \dots & k(x_n, x_n) \end{bmatrix} \right).$$

Definition 4.3 (Mean and Covariance Function). *We define the mean function and covariance kernels as*

$$m(x_i) := \mathbb{E}[f(x_i)]$$

and

$$k(x_i, x_{i'}) := \text{cov}(f(x_i), f(x_{i'})).$$

Some common examples of Gaussian processes include

- Brownian motion on \mathbb{R} :

$$m \equiv 0,$$

and

$$k(s, t) = \min(s, t)$$

- Ornstein Uhlenbeck process with parameters θ and σ :

$$m \equiv 0,$$

and

$$k(s, t) = \frac{\sigma^2}{2\theta} \left(e^{-\theta|t-s|} - e^{-\theta(t+s)} \right)$$

We can generate novel Gaussian processes with a custom m , and covariance kernel k . For the sake of this thesis we only consider symmetric k that are also stationary - that is $k(x, x')$ can be expressed as a function of $x - x'$.

Definition 4.4 (Positive Semi-Definite Matrix). *An $n \times n$ matrix \mathbf{A} is positive semi-definite if $\mathbf{v}^T \mathbf{A} \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$.*

Theorem 4.1 (Sufficient Condition for Positive Semi-Definite). *A symmetric matrix \mathbf{A} is positive semi-definite, if (and only if) it's eigenvalues are non-negative.*

Proof.

□

Definition 4.5 (Positive Semi-Definite Kernel). *A kernel $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is positive semi-definite if the matrix*

$$\mathbf{K} = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ k(x_n, x_1) & \dots & \dots & k(x_n, x_n) \end{bmatrix}$$

is positive semi-definite for any collection of $x_i \in \mathcal{X}$

Theorem 4.2. *All symmetric positive semi-definite matrices are covariance matrices for some set of random variables*

Proof.

□

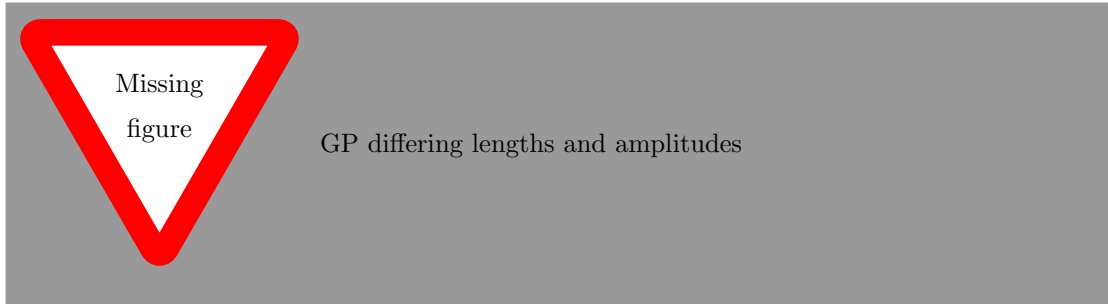
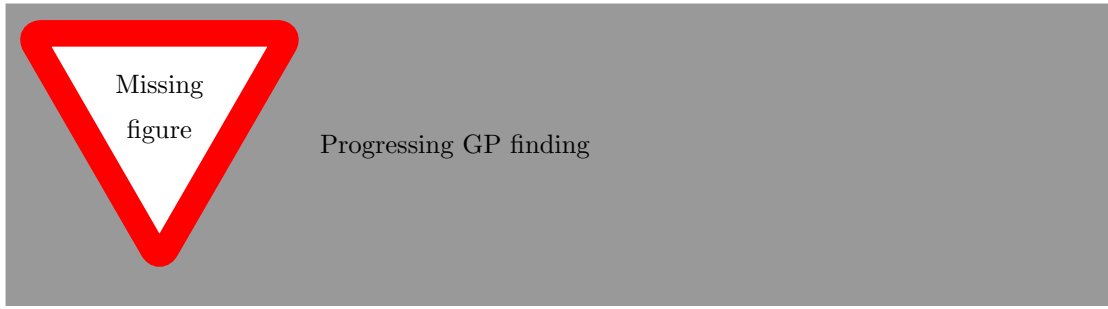
Some common

Theorem 4.3 (Positive Semi-Definiteness of RBF). *The radial basis function $\frac{1}{\sigma^2} \exp(-\frac{(x-x')^2}{2\sigma^2})$ is a positive semi-definite kernel.*

Theorem 4.4 (Bochner's Theorem). *Let k be a stationary kernel function such that $k(x, x') = f(d)$. A function $k : \mathbb{R}^d \rightarrow \mathbb{C}$ is the covariance function of a weakly stationary mean square continuous complex-valued random process of \mathbb{R}^d if and only if it can be represented as*

$$k(\tau) = \int_{\mathbb{R}^d} \exp(2\pi i \mathbf{s} \cdot \tau)$$

(Rasmussen and Williams 2008, p. 82)



Adding in observation variance

4.2 Kernel Families

Choice of kernel largely depends on the nature of the function to be estimated. Smooth kernels such as the squared exponential kernel are favourable when estimating functions that are smooth, since it is infinitely mean square differentiable.

Squared Exponential

The squared exponential kernel is of the form

$$k(x, x') = \sigma^2 \exp \left(-\frac{\|x - x'\|^2}{2\ell^2} \right)$$

Matérn Class

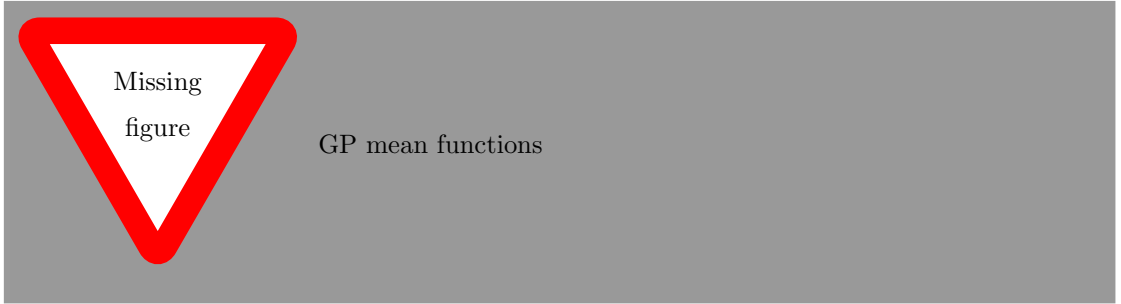
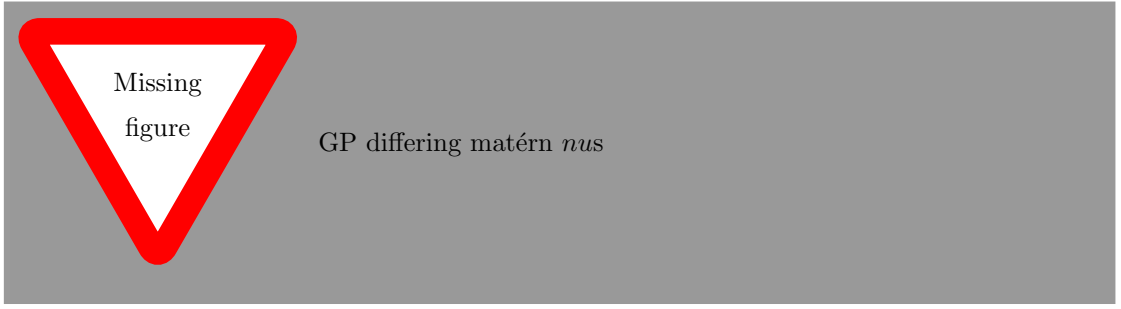
The Matérn class allows you to choose how smooth you think the function is.

The Matérn exponential kernel is of the form

$$k_\nu(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\|x - x'\|}{\ell} \right)^\nu K_\nu \left(-\frac{\sqrt{2\nu}\|x - x'\|}{\ell} \right)$$

where K_ν is a modified Bessel function (see Abramowitz and Stegun 2013, p. 374). The general form is not very insightful, but for the most common values of $\nu = 1/2, 3/2$ and $5/2$, we have the forms:

$$k_{1/2}(x, x') = \sigma^2 \exp \left(-\frac{\|x - x'\|}{\ell} \right)$$



$$k_{3/2}(x, x') = \sigma^2 \left(1 + \frac{\sqrt{3} \|x - x'\|}{\ell} \right) \exp \left(-\frac{\sqrt{3} \|x - x'\|}{\ell} \right)$$

$$k_{5/2}(x, x') = \sigma^2 \left(1 + \frac{\sqrt{5} \|x - x'\|}{\ell} + \frac{5 \|x - x'\|^2}{3\ell^2} \right) \exp \left(-\frac{\|x - x'\|^2}{2 * \ell^2} \right)$$

Zero mean processes with a Matérn kernel are n times mean square differentiable, for all $n < \nu$.

$\nu = 1/2$ is an Ornstein-Uhlenbeck process for one dimension.

It asymptotically approaches the squared exponential class.

4.3 Differing mean functions

Chapter 5

Stuff that could fix stuff

- Gaussian process has gaussian process for variance, run any set of parameters multiple times, have very large length scale, maybe a prior on it
- Add in an artificial variance for the values where the disease dies off
- t-process to allow for sudden drop off at threshold
- Maybe use matern to allow for sudden jump but I think that allows for very small sudden jumps
- Used a relative distance squared

Bibliography

- Abramowitz, Milton and Irene A. Stegun, eds. (2013). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. 9. Dover print.; [Nachdr. der Ausg. von 1972]. Dover books on mathematics. New York, NY: Dover Publ. 1046 pp. ISBN: 978-0-486-61272-0.
- Adams, John H. and Ivo Mueller (Sept. 2017). “The Biology of *Plasmodium vivax*”. In: *Cold Spring Harbor Perspectives in Medicine* 7.9, a025585. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a025585. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5580510/> (visited on 03/24/2023).
- Aron, Joan L. and Robert M. May (1982). “The population dynamics of malaria”. In: *The Population Dynamics of Infectious Diseases: Theory and Applications*. Ed. by Roy M. Anderson. Boston, MA: Springer US, pp. 139–179. ISBN: 978-1-4899-2901-3. DOI: 10.1007/978-1-4899-2901-3_5. URL: https://doi.org/10.1007/978-1-4899-2901-3_5.
- Champagne, Clara et al. (Jan. 2022). “Using observed incidence to calibrate the transmission level of a mathematical model for *Plasmodium vivax* dynamics including case management and importation”. In: *Mathematical Biosciences* 343, p. 108750. ISSN: 00255564. DOI: 10.1016/j.mbs.2021.108750. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0025556421001541> (visited on 08/22/2023).
- Cowman, Alan F. et al. (2016). “Malaria: Biology and Disease”. In: *Cell* 167.3. Type: Review, pp. 610–624. DOI: 10.1016/j.cell.2016.07.055. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84994000411&doi=10.1016%2fj.cell.2016.07.055&partnerID=40&md5=81d9b4c51fe738ac66e0c8561b12c5bf>.
- Hagenaars, T. J., C. A. Donnelly, and N. M. Ferguson (Apr. 2006). “Epidemiological analysis of data for scrapie in Great Britain”. en. In: *Epidemiology and Infection* 134.2, pp. 359–367. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268805004966. URL: https://www.cambridge.org/core/product/identifier/S0950268805004966/type/journal_article (visited on 03/26/2024).
- Keeling, Matthew James and Pejman Rohani (2008). *Modeling infectious diseases in humans and animals*. OCLC: ocn163616681. Princeton: Princeton University Press. 366 pp. ISBN: 978-0-691-11617-4.
- Milner, Danny A. (Jan. 2018). “Malaria Pathogenesis”. en. In: *Cold Spring Harbor Perspectives in Medicine* 8.1, a025569. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a025569. URL: <http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a025569> (visited on 03/24/2023).
- Price, R.N. et al. (2020). “*Plasmodium vivax* in the Era of the Shrinking *P. falciparum* Map”. English. In: *Trends in Parasitology* 36.6, pp. 560–570. ISSN: 1471-4922. DOI: 10.1016/j.pt.2020.03.009.
- Rasmussen, Carl Edward and Christopher K. I. Williams (2008). *Gaussian processes for machine learning*. 3. print. Adaptive computation and machine learning. Cambridge, Mass.: MIT Press. 248 pp. ISBN: 978-0-262-18253-9.

- Smith, David L. et al. (Apr. 2012). “Ross, Macdonald, and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens”. en. In: *PLOS Pathogens* 8.4. Publisher: Public Library of Science, e1002588. ISSN: 1553-7374. DOI: 10.1371/journal.ppat.1002588. URL: <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1002588> (visited on 03/28/2023).
- White, Michael T. et al. (Mar. 30, 2016). “Variation in relapse frequency and the transmission potential of *Plasmodium vivax* malaria”. In: *Proceedings of the Royal Society B: Biological Sciences* 283.1827, p. 20160048. ISSN: 0962-8452, 1471-2954. DOI: 10.1098/rspb.2016.0048. URL: <https://royalsocietypublishing.org/doi/10.1098/rspb.2016.0048> (visited on 08/22/2023).
- World Health Organization (Dec. 2022). *World malaria report 2022*. en. Tech. rep. Geneva: World Health Organization.
- Zekar, Lara and Tariq Sharman (2023). “Plasmodium Falciparum Malaria”. eng. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing. URL: <http://www.ncbi.nlm.nih.gov/books/NBK555962/> (visited on 03/24/2023).
- Zha, Wen-ting et al. (2020). “Research about the optimal strategies for prevention and control of varicella outbreak in a school in a central city of China: based on an SEIR dynamic model”. en. In: *Epidemiology and Infection* 148, e56. ISSN: 0950-2688, 1469-4409. DOI: 10.1017/S0950268819002188. URL: https://www.cambridge.org/core/product/identifier/S0950268819002188/type/journal_article (visited on 03/26/2024).