Stockholms universitet

L3, Timing and observations + Endemic models

Tom Britton

June 2020

Stockholms universitet

The basic reproduction number

Recall: R_0 = expected number individuals a typical infected person infects when everyone is susceptible

R_0 depends both on disease (infectious agent) and on community!!

$R_0 < 1$ or $R_0 > 1$ makes a very big difference!

Next page: R_0 for some diseases (and communities and time periods), Anderson and May, 1991

Stockholms universitet

R_0 for some diseases, communities and time periods (Anderson & May, 1991)

Table 4.1 Estimated values of the basic reproductive rate, R_0, for various infections (data from Anderson (1982*b*), Anderson and May (1982*d*, 1985*c*, 1988), Anderson *et al.* (1988), Nokes and Anderson (1988)).

Infection	Geographical location	Time period	R_0
Measles	Cirencester, England	1947–50	13–14
	England and Wales	1950–68	16–18
	Kansas, USA	1918–21	5–6
	Ontario, Canada	1912–13	11–12
	Willesden, England	1912–13	11–12
	Ghana	1960–8	14–15
	Eastern Nigeria	1960–8	16–17
Pertussis	England and Wales	1944–78	16–18
	Maryland, USA	1943	16–17
	Ontario, Canada	1912–13	10–11
Chicken pox	Maryland, USA	1913–17	7–8
	New Jersey, USA	1912–21	7–8
	Baltimore, USA	1943	10–11
	England and Wales	1944–68	10–12
Diphtheria	New York, USA	1918–19	4–5
	Maryland, USA	1908–17	4–5
Scarlet fever	Maryland, USA	1908–17	7–8
	New York, USA	1918–19	5–6
	Pennsylvania, USA	1910–16	6–7
Mumps	Baltimore, USA	1943	7–8
	England and Wales	1960–80	11–14
	Netherlands	1970–80	11–14
Rubella	England and Wales	1960–70	6–7
	West Germany	1970–7	6–7
	Czechoslovakia	1970–7	8–9
	Poland	1970–7	11–12
	Gambia	1976	15–16
Poliomyelitis	USA	1955	5–6
	Netherlands	1960	6–7
Human Immunodeficiency Virus (Type I)	England and Wales (male homosexuals)	1981–5	2–5
	Nairobi, Kenya (female prostitutes)	1981–5	11–12

Exercise 3.1: Why is $R_0 > 1$ for all diseases above?

Stockholms universitet

Time events and terminolgy

Infection-related events in time

Infection time: calendar time when someone gets infected (unobserved)

Report time: calendar time an infected is reported a case (observed for reported cases)

Onset of symptoms: calendar time when an infected notices symptoms (often observed for reported cases)

Incubation period: relative time between getting infected and having symptoms (unobserved)

Latent period: relative time between getting infected and becomen infectious (SEIR models) (unobserved)

Generation time and serial interval: relates times between infector and infectee – see next slides

Stockholms universitet

Generation time (Britton and Scalia Tomba, 2019)

Most important property, probably R_0: quantifies how many new infections (on average) infected people cause in the beginning of an epidemic outbreak

Second most important, if we are interested in time evolution (or if we want to estimate R_0 from reported incidence over time!): the **generation time** G

G is the time between getting infected and infecting a new person

$\Longrightarrow$ individuals who infect more than one individual generate several generation times (and individuals who infects noone generate no generation times)

Generation times are not all the same, so G is a **random variable**

$g(s)$ denotes the **generation time distribution** for G. E.g. $g(s)$ is Normal or *gamma distribution* with a given mean and st.d.

Stockholms universitet

Toy example

Suppose that $R_0 = 2$, and each infected infects one individual after 1 week and one individual after 2 weeks ($g(1) = g(2) = 0.5$)

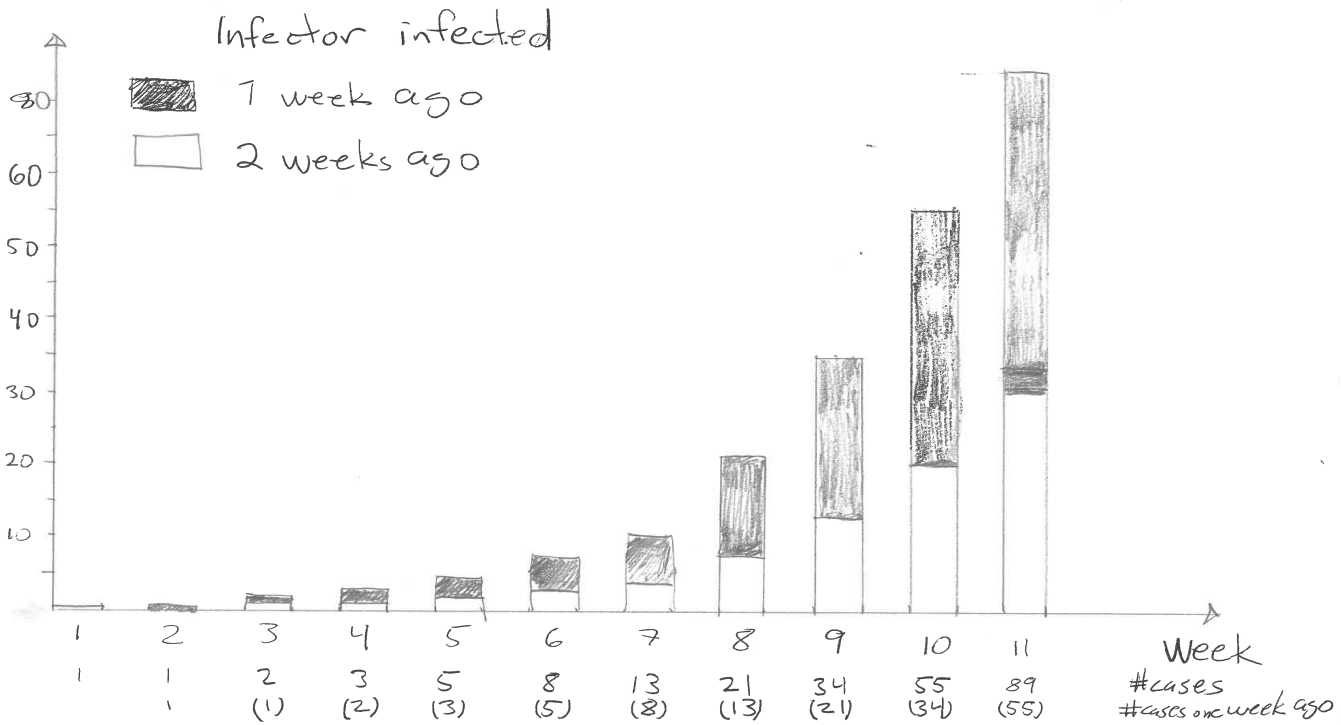
What is $E(G)$?

Stockholms universitet

Toy example

Suppose that $R_0 = 2$, and each infected infects one individual after 1 week and one individual after 2 weeks ($g(1) = g(2) = 0.5$)

What is $E(G)$? 1.5 weeks, and $st.d.(G)$? 0.5 weeks (below plot of # infections each week)



Stockholms universitet

Estimating $g(s)$

How estimate generation time distribution $g(s)$?

Answer: **Contact tracing**: For some identified cases, it is traced by whom and when they were infected

This gives some observed generation times $g_1, \ldots, g_k$

This is often only way, but problematic:

- Generation time defined forward in time but contact tracing backward in time. Problematic?
- For some cases a unique infector and infection time is identified, but for some there are several possibilities (and some have none)
- onset of symptoms more common to observe than infection times
- Identified cases are often severe cases. Do mild/asymptomatic cases have same generation times?

Looking backwards: contact tracing

Fibonacci numbers and the Golden ratio ...

$\Longrightarrow$ The mean generation time when contact tracing will be < 1.5

So if you estimate $E(G)$ (or all of G) from contact tracing you will *under-estimate* $E(G)$

(Next lecture: implies you will also *under-estimate* R_0!!)

Stockholms universitet

Generation times vs Serial intervals

Serial intervals instead of generation times

(We now forgetproblem of looking backwards)

Infection times are hardly ever observed, but onset of symptoms are

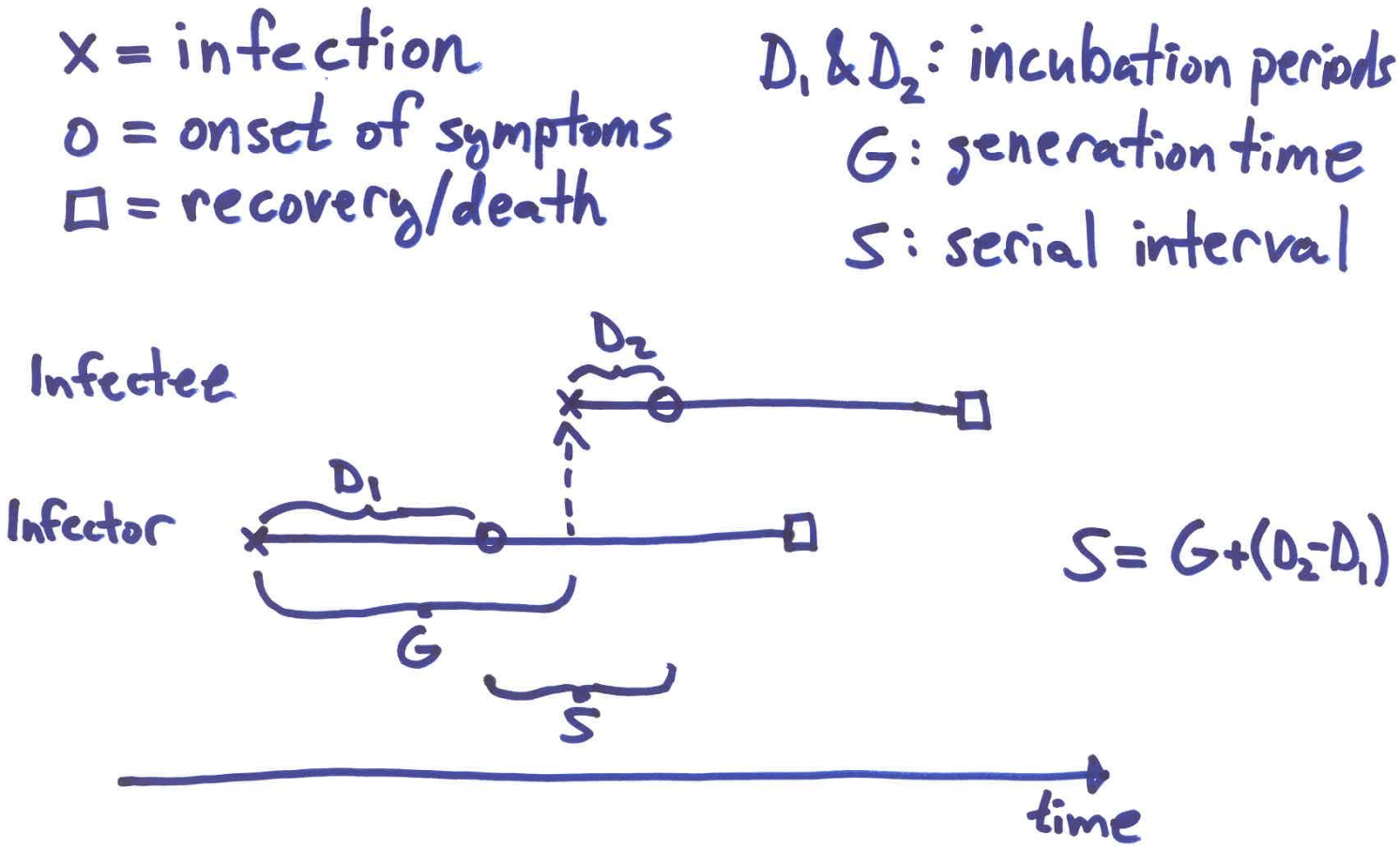
G = time between infection times (unobserved)

S = time between onset of symptoms (observed)

Stockholms universitet

Generation times vs Serial intervals, cont'd

Generaton times vs Serial intervals



Stockholms universitet

Generation times vs Serial intervals, cont'd

$\Longrightarrow S = G + (D_2 - D_1)$ (D_1 and D_2 = incubation periods of infector and infectee)

So, if incubation times are independent and independent of G, then

$E(S) = E(G)$, and $V(S) \geq V(G)$

(The relation holds true for all (?) epidemic models)

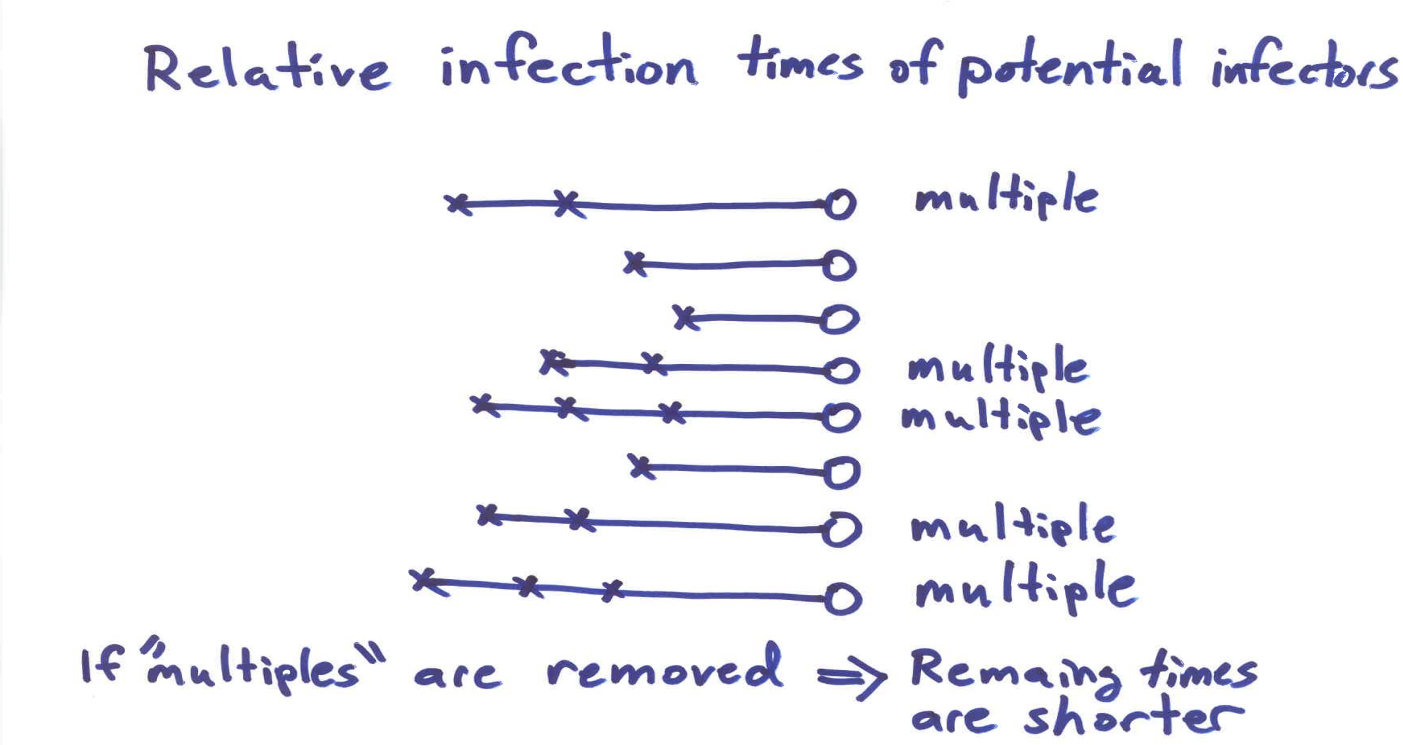
So, if we estimate $G \sim \{g(s)\}$ from observations on Serial intervals we will *over-predict* variance of G

(Next lecture: implies you will also *under-estimate* R_0!!)

Stockholms universitet

Multiple exposures

Another problem when contact tracing is that sometimes there are several potential infectors (see illustration on next slide)



Stockholms universitet

Biases for Ebola and Corona

If observations with more than one infected are neglected, remaining intervals are biased from below.

This will also lead to *under-estimation* of $E(G)$ (Next lecture: and of R_0!!)

Bias effect for Ebola and Covid-19

For Ebola 75% of contacts had multiple potential infectors. The combinded under-estimation of R_0 was $\approx 23\%$

For Corona (Covid19) there was no information of multiple infectors (but I am sure there were!), so only considering bias from backward tracing we believe R_0 is under-estimated by $\approx 12\%$.

Stockholms universitet

Variation in response + under-reporting

Many infectious disease have variable response: some infected have very severe symptoms or may even die, others have milder symptoms, and some infected show no symptoms but may still infect others

$\Longrightarrow$ **Severe, mild, asymtomatic cases**

Observed cases have bias towards severe cases

Can in turn lead to biases in estimates of infection falatity risk (ifr) – compare with case fatality risk (cfr)

For nearly all infectious diseases the number of reported cases is small in comparison the number of infected: **under-reporting**

One way to overcome bias in symptoms and bias in fraction infected: test **random samples** in community for virus and/or antibodies

Stockholms universitet

Possible tasks for individual project 2

Project 2 is done individually. One option is to investigate some of the feastures mentioned above numerically by simulations. A good advice is to look in Britton and Scalia Tomba for more details (quite technical)

How big is difference between true generation distribution and estimated from contact tracing (looking backwards) of some individuals?

For some model for single/multiple infections, how big is the effect of throwing away cases with multiple infections

What if there are symptomatic and asymptomatic where a certain fraction become symptomatic and the rest asymptomatic, symptomatic infect more new cases, there generation time differs a bit, and only generastion times from symptomatic-symptomatic infections are observed?

Stockholms universitet

Endemic diseases

When interest is on long-term situation (as opposed to short term outbreaks) the assumption of a fixed population must be relaxed

Consider an SIR disease in a population where individuals die and new are born. Assume:

- SIR disease (life long immunity)
- population at "equilibrium" (in terms of size and incidence)
- disease endemic (constantly present, no big fluctuations)
- $\tilde{s}$, $\tilde{i}$ and $\tilde{r}$ denote the average fractions susceptible, infectious and removed
- R_0 = average number of infections caused by one individual – if everyone was susceptible!

Think of childhood diseases (e.g. chicken-pox)

Stockholms universitet

Endemic diseases, expression for $\tilde{s}$

When disease is in endemic equilibrium each infected individual on average infects exactly 1 new person!

Given R_0 and $\tilde{s}$ an infected individual infects on average $R_0\tilde{s}$ new individuals

Stockholms universitet

Endemic diseases, expression for $\tilde{s}$

When disease is in endemic equilibrium each infected individual on average infects exactly 1 new person!

Given R_0 and $\tilde{s}$ an infected individual infects on average $R_0\tilde{s}$ new individuals

$\Longrightarrow R_0\tilde{s} = 1$!!

$\tilde{s} = \frac{1}{R_0}$

$\tilde{s}$ = average fraction susceptible = $\frac{\text{average age at infection}}{\text{average life-length}}$

Exercise 3.2 Suppose $R_0 = 1.5$, 3 and 6 respectively, compute $\tilde{s}$.

Exercise 3.3 Suppose you know that average age of infection of chickenpox is 8 years, and average life-length is 80 years. Estimate R_0 for chickenpox.

Endemic diseases, expression for $\tilde{i}$

If ι is the average length of infectious period and ℓ average life-length, then ι/ℓ is the average time of the life an individual is infectious

Since population/disease in equilibrium this is also the population fraction of infectives

$$\tilde{i} = \frac{\iota}{\ell}$$

Stockholms universitet

Exercises

Exercise 3.4 Consider an endemic disease with one week infectious period and a population with 75 years expected life-length. Compute the average fraction infective $\tilde{i}$.

Exercise 3.4 Consider the disease in the previous exercise and consider the Icelandic population ($n = 250\ 000$). What is the average *number* of infectives? How about England ($n = 60\ 000\ 000$)?

Exercise 3.6 What do you think will happen with the disease in the two countries (remember that if the number of infectives drops to 0 the disease goes extinct - until it is "re-imported")?