

Generation times in epidemic models

Gianpaolo Scalia Tomba

Dept Mathematics, Univ of Rome "Tor Vergata", Italy

in collaboration with

Åke Svensson, Dept Mathematics, Stockholm University, Sweden

Tommi Asikainen and Johan Giesecke, ECDC, Sweden

InFER 2011, Warwick 28/3 - 1/4 2011

In the infectious disease setting, generation time refers to the length of the time interval between one individual's time of infection and the time when this individual infects another individual, in analogy with the use of the term in demography.

Some early users of the concept are Hope-Simpson and Bailey, in the '50s and '60s, in relation to diseases like measles. In measles, the latency period is long and the infectious period short and, in localized outbreaks, it is usually possible to clearly distinguish the first two or three "generations" of infections.

Recently, the concept of generation time has been used in formulae relating initial spread of an infection to its reproduction number, following the classical demographic definitions [see eg Keyfitz & Caswell, 2005]

$l(a) := P(\text{infectious period is longer than } a)$

$m(a) := \text{intensity of infectious contact at infection age } a$

=> Malthusian parameter, as the unique value of $r \geq 0$ satisfying

$$1 = \int_0^{\infty} e^{-ra} l(a) m(a) da$$

We also have $R_0 = \text{expected total offspring of an individual} =$

$$\int_0^{\infty} l(a) m(a) da$$

We then define the "cohort generation time distribution" as

$$g(a) = \frac{l(a)m(a)}{\int l(u)m(u)du}$$

and the "average cohort generation time" as the expected value in this distribution, i.e.

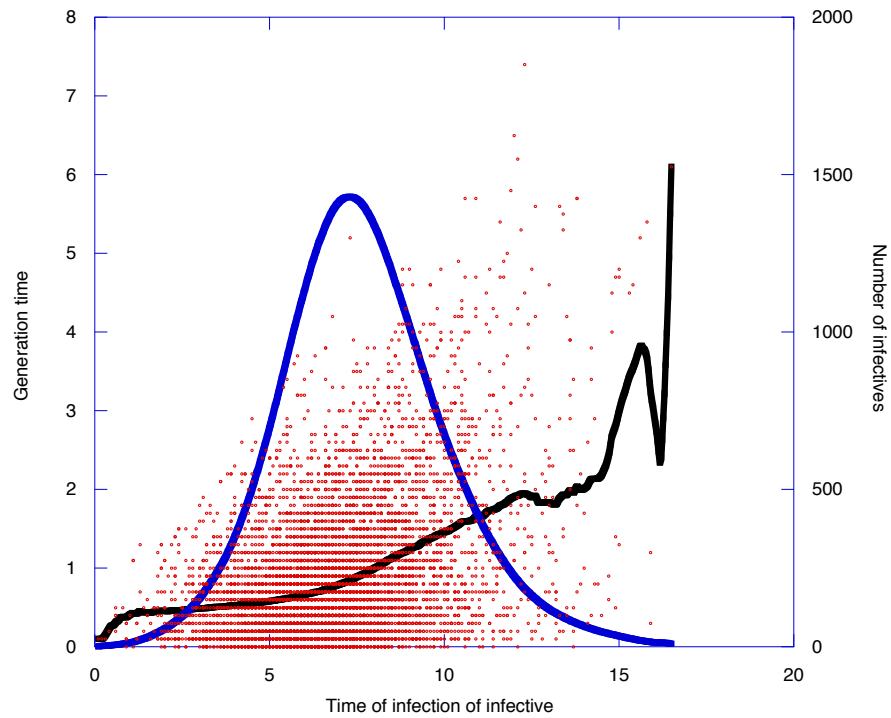
$$T_c = \int_0^{\infty} ag(a)da$$

All these relations can e.g. be expressed in terms of the Laplace transform of the density g . By linearization in the origin we can for instance approximate:

$$R_0 \approx 1 + rT_c$$

Problem:

in a simulation of a simple stochastic epidemic, with generation times measured "backwards", from each infection to the time of infection of the infector, the following picture of the time evolution of average GT came out:



If the generation time "changes" during the epidemic, what does the theoretical relationship between r and R_0 refer to?

Findings...

There is a mixture of problems of definition, effects of sampling schemes and effects of the epidemic dynamics...

- "hidden statistical implication of the demographic definitions"
- difference between "forward" and "backward" observation
- how to compute averages in the forward case
- epidemic specific dynamics in the forward and backwards cases

Demographic tradition counts daughters to mothers and e.g. evaluates the age of the mother at each time a daughter is born -> average age at childbearing in the cohort generation time distribution.

However, in statistical terms, this amounts to "size-biased sampling" of mothers, since each mother will be present in the sample with a multiplicity equal to her number of daughters.

Example:

suppose we follow two individuals (A and B, say) from their respective moments of infection and observe (IP= infectious period, GT = generation time):

- A: IP=4 and secondary cases at times 1,2,3 -> av. GT = 2

- B: IP=8 and secondary cases at times in 1,2,3,4,5,6,7 -> av. GT = 4

Should we calculate the general average GT as 3 (infector based) or $34/10$ (infectee based)?

At this point, it is convenient to define a mathematical model for the infection process, in which to carry out the analysis of properties of generation times and observations. The "simplest" model is the so called SIR model:

Population of fixed size N at time $t=0$, 3 classes $S(t)+I(t)+R(t) = N$, $S(0)=N-1$ & $I(0)=1$.

The "mass action" principle = homogeneous mixing ->
the number of new infections in a time interval $[t,t+dt]$ is expected to be
 $\beta S(t)I(t) dt$

Usually leads to ODE system ($\mu = 1/\text{average length of infectious period IP}$)

$$S' = -\beta SI$$

$$I' = \beta SI - \mu I$$

$$R' = \mu I$$

Another way of viewing the situation is that each person has a given average number of contacts/time unit and that only those of an I person with an S lead to infections with a given probability.
Usually leads to Markov process description

$(S, I, R) \rightarrow (S-1, I+1, R)$ with intensity $\alpha I S / N$

$(S, I, R) \rightarrow (S, I-1, R+1)$ with intensity μI

In an average sense, $\beta = \alpha / N$, but only if N remains constant and R people remain in the contact population...

Summary of behaviour

-Threshold $R = \alpha/\mu =$ expected number of secondary cases when all contacts are susceptible. Behaviour of spread radically different in large populations below or above $R = 1$.

-Initial exponential growth with "speed" $\alpha-\mu$ (ODE->linearization, Markov process -> branching process approximation incl Pr(extinction) although supercritical)

-If epidemic, stochastic trajectory close to deterministic trajectory = unimodal infectives curve, slightly quicker at start than at end...

- Final size: if epidemic, the proportion τ satisfying $\tau = 1 - \exp(-R_0\tau)$, will be infected. τ is less than 1...

The most useful of these results: the Herd Immunity effect

since $R \leq 1$ leads to early extinction (of disease !), immunization of proportion $1 - 1/R$ of population is sufficient to avoid epidemics...

Current research on models is focused on heterogeneity in contacts, e.g. by introducing "households" or "networks" in the contact specifications...

Small interlude about Poisson processes and calculations with Exp...

An I individual is infective during an infectious period T with $\text{Exp}(\mu)$ distribution. During this period, contacts are made with constant intensity α and these lead to infection with a certain probability $(S(t)/N)$ dependent on time (thinning).

Then: the longer T , the larger the number of infected & viceversa...

Furthermore, the time points of infection in the interval are the generation times related to the considered infective individual...

If the thinning probability σ is essentially constant in the interval, the events will be scattered as iid uniform on $[0, T]$.

-> average position of an event $E(T)/2$ (actually slightly more, because of need to condition on "at least one event" -> $E(T | \# \text{events} \geq 1) = 1/\mu + 1/(\mu + \alpha\sigma)$)

One can thus equally reasonably define

- "natural history based generation time"

(average of infector's ages at secondary cases)

=> average GT is slightly larger than half the average IP

- "infectee based generation time"

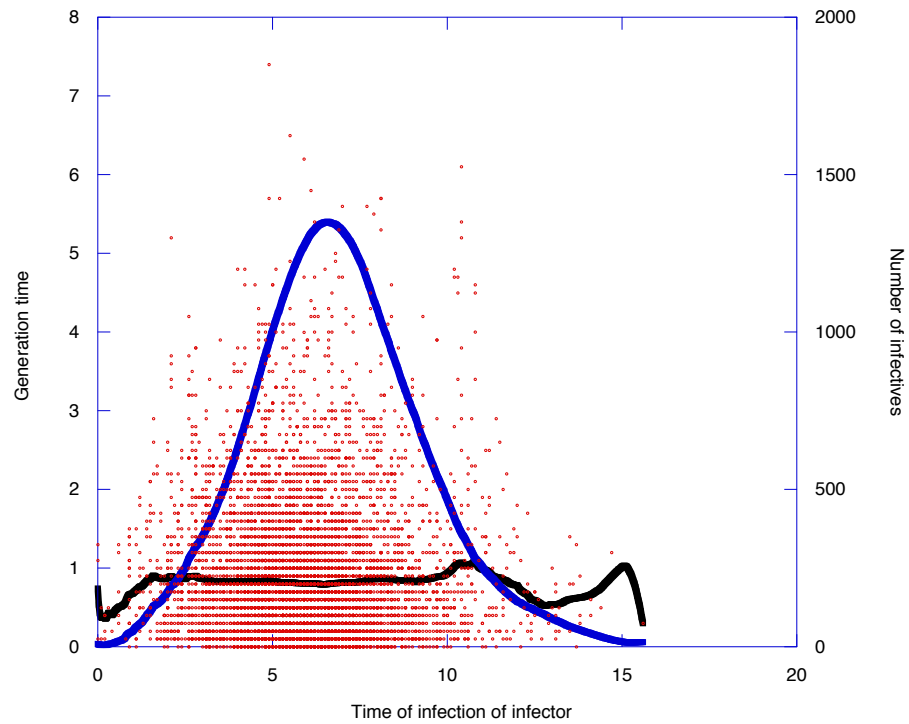
(the demographic definition used above)

=> average GT equals the average IP (in the Markov case...)

However, what we define to be a "generation time" doesn't really matter. What matters, from a statistical point of view, is that what we observe corresponds to what we assume that we observe...

It turns out that there is a difference between what can be called "forward" observation (following an infected person and registering his secondary cases, if any) and "backward" observation (starting from an infected person and finding the time of infection of his infector).

If a complete epidemic is simulated and all infectious contacts and corresponding times are recorded, various observation schemes can be reconstructed...



Generation times (dots) plotted as single points at time of infection of infector.

Expected behaviour: size biased sampling -> average ≈ 1 (same as average IP...). based on 7888 infectees.

This picture gives approximately the result expected from the demographic theory (average 1 with used parameters).

It should be noted that each infector participates in the data set as many times as he has produced secondary cases. If the average generation time is first computed "within" infector and then plotted, size-bias is avoided and a lower average level is obtained (average about $2/3$).

However, if we observe GTs "backwards", from each infectee back to the infector, we get a completely different behaviour (first graph shown), we get to see the "infection age distribution of infectors", at least in models that use equal infection rates for all potential infectors...

If the number of potential infectors is growing exponentially, as e.g. in the first part of an epidemic, there will be an "overrepresentation" of recently infected individuals => shorter average GT than "expected".

As the epidemic proceeds, the average "infection age" increases and so do the average generation times (average from 1/2 to about 2...).

From a statistical point of view, it is interesting to note that the "set of generation times" is the same in the three different calculations above...

The population dynamics of the epidemic play a role...

In the backwards situation, the exponential growth rate of new infectives decreases all the time, due to the progressive depletion of susceptibles => average GT becomes longer and longer...

In forward measurements, independently of size biased sampling or not, GTs will, on average, become shorter as the rate of depletion of susceptibles becomes large (maximum depletion rate slightly before peak of infectives), because the average time of infection during an infector's IP has uniform distribution if the infection rate is constant, otherwise inherits the shape of the changing infection rate: decreasing infection rate => more probability of secondary cases in beginning of IP compared to end... (assuming a model with constant meeting rate times prob of meeting a susceptible...).

Conclusions

The presented results, although model dependent, probably represent generic problems related to the analysis of generation times.

Just like R_0 , the generation time distribution is a "constructed" parameter, not a descriptor of disease natural history, and therefore not necessarily constant in all circumstances.

Its original appeal and use is based on clearly identifiable chains of "parents" and "offspring". This is not so clear in, say, an influenza epidemic...

One could try to "use" the described mechanisms for statistical purposes.

For instance, one often described use of the relation between generation times and initial exponential growth rate is the estimation of R_0 from the first phase of an epidemic.

In the simple Markov SIR model, where $R_0 = \beta/\mu$, the average generation time T is expected to be $1/\mu$ and the growth rate $r = \beta - \mu$, which leads to the formula $R_0 \approx 1 + rT$.

However, if backward observation = infector tracing is performed in the initial phase of an epidemic, the expected average generation time $T = 1/\beta$ and the formula should now be $R_0 \approx 1/(1 - rT)$.

In the discussion about the statistics of generation times, one should not forget that the moments of infection are generally not observable and that only surrogates, e.g. serial times = time interval between symptoms, may be observable. Also, the question about "who infected whom" is not always easy to answer...

Since most models assume infectious periods and contact rates as basic and constant ingredients, it is probably safer to seek to estimate these quantities, after proper consideration of the sampling mechanism of observations, than to think in terms of "generation times".