Mathematical Statistics
Stockholm University

A note on generation times in epidemic models

Åke Svensson,
The Swedish Institute for Infectious Disease Control
and Stockholm University

Research Report 2005:14

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se/matstat

Mathematical Statistics
Stockholm University
Research Report **2005:14**,
http://www.math.su.se/matstat

A note on generation times in epidemic models

Åke Svensson,
The Swedish Institute for Infectious Disease Control
and Stockholm University*

December 2005

Abstract

Generation times measures the time between the infection of a primary case and one of its secondary cases. The distribution (and mean) of the generation time is derived for a rather general class of epidemic models. The relation to assumptions on distributions of latency times and infectious times or more generally on random time varying infectiousness, is investigated. Serial times, defined as the time between occurrence of observable events in the progress of an infectious disease (e.g. the onset of clinical symptoms), are also considered.

Key words: epidemic models, generation times, serial times, transmission intervals

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: akes@math.su.se.

A note on generation times in epidemic models

Åke Svensson

Mathematical Statistics, Stockholm University and the Swedish Institute for Infectious Disease Control

December 19, 2005

Abstract

Generation times measures the time between the infection of a primary case and one of its secondary cases. The distribution (and mean) of the generation time is derived for a rather general class of epidemic models. The relation to assumptions on distributions of latency times and infectious times or more generally on random time varying infectiousness, is investigated. Serial times, defined as the time between occurrence of observable events in the progress of an infectious disease (e.g. the onset of clinical symptoms), are also considered.

Keywords: epidemic models, generation times, serial times, transmission intervals

1. Introduction

The reason for writing this short note is to try to provide some understanding of statistical and stochastic properties related to generation (and serial) times as used in models for spread of communicable diseases.

Traditionally a main input to epidemic models is assumptions on the distributions of individual latency times, infectious times and infectiousness (e.g. represented as the virus shedding process). These concepts are all related to the the progress of an infection within an infected individual. They describe the natural history of an infection and should, in principle, be observable from studies of single infected persons. We will call these aspects of the infection inter-individual. Medical handbooks, that are focused on clinical aspects of infections and infectious diseases provide some information about these.

Further to inter-individual aspects of infections an epidemic model also has to describe how, why and when infections are transmitted between individuals. Assumptions about transmission explain why epidemic processes may behave differently in populations with different contact patterns. Typically structures as households and working places have to be taken into account.

Lately the notion of generation times has been popular and now is an important concept in models of the SARS epidemic and pandemic influenza (Wallinga and Teunis (2004) and Ferguson et al. (2005)). Generations times are assumes to reflect the time between the infections of a primary its secondary cases. A general discussion of the concept which includes a survey of several infections is given by Fine (2003). Fine uses the term transmission interval.

Generation times are related to the duration of incubation and infectious periods as well as time-variation in infectivity that an infected person emits. By definition a generation time involves a primary case and a secondary case infected by a primary case. Thus, it can only be understood in relation to a model for transmission between individuals.

One way to give generation time a well-defined, and general, meaning is to consider its meaning in the most simple conceivable model. This would be the model where infectious contacts are taken according to a homogeneously mixing assumption, i.e., an infected persons is equally likely to contact any other member of the population. An alternative way is to study generation times in more complex models. In that case the mean and variability of generation times will have to be considered together with the laws that governs the transmission in the particular complex model.

The discussion in this paper will cover both these alternatives. The derivation will be based on a random measure K, that describes how infectiousness of an infected person changes over time (see section 3). Infected persons may have different potentiality of spreading the infections. In the simple model, with homogeneous mixing, the variability of K will only involve differences between the inter-individual progress of an infection within infected individuals. In more complex models it will also involve differences in how individuals spread the infection to other persons.

The possibility to transmit an infection will change during a epidemic, due to the possibility that already infected persons may have obtained immunity. This may influence generation times, in particular in the intense period of epidemic spread. For this reason we will restrain the analysis to the beginning of an epidemic (or to situations where an infected individuals infectivity is not influenced by changes in immunity in the total population).

The following discussion is based on a general model for infectious spread (section 3). In section 4 result on distributions and moments for generation times are derived. These results are applied to some special models in section 2. Some remarks regarding the possibility to observe generations times in studies of real epidemic outbreaks are made in section 5. Finally some comments of the more general concept of serial times are given in section 7.

Much of the discussion that follows is based on well-known results from the theories of demography (Keyfitz and Caswell (2005)) and branching processes (Jagers (1975), Kimmel and Axelrod (2002) and Haccou et al. (2005)). For formal proofs of technical results we will refer to these texts. We will leave out standard derivations of distribution functions and means.

2. Generation time

Definitions of generation time formulated in literature, are in most cases essentially verbal. They may look common-sense and straight-forward, but it is not obvious how to translate them to mathematical well-defined concepts.

We quote some definitions:

- "*the sum of the average latent and the average infectious period is referred to as the average generation time of the infection*". Anderson and May (1992).

- "*the time between the appearance of similar symptoms (e.g. rash, coughs) in successive generations*" Giesecke (2002).
- "*the generation interval is the time from symptoms onset in a primary case to symptom onset in a secondary case*" Wallinga and Teunis (2004).
- "*The average time taken for the secondary cases to be infected by a primary case*" Anderson et al. (2004).

Some of these definitions are made in the context of particular assumption of duration of latency and infectious times and may not be immediately translatable to other models. With the exception of the definition of Anderson and May (1992) none of these definitions are stated in mathematical terms. Since any reasonable epidemic model involves random elements due to both variation between individuals and to randomness in transmission of the infectious agent, it is necessary to consider the stochastic properties of generation times. We may want to find the distribution of a "typical" generation time or at least the mean of a "typical" generation time.

It should also be observed that two of the quoted definitions (Anderson and May (1992) and Anderson et al. (2004)) relate to times between successive infections and the other (Giesecke (2002) and Wallinga and Teunis (2004)) refer to any identifiable event in the progress of an infection. This is a more general definition and we will here refer to such intervals as serial times. From a statistical viewpoint there are important differences between generation and serial times. Loosely speaking, generation time refers to what happens to a single infected individual, whereas serial time is a mixture what happens in two infected individuals. Some consequences of this are discussed in section 7.

One of the difficulties is that in most models (and most real applications) there are infected person that causes no secondary cases. It seems natural to define the the generation time to be infinite. This will make it impossible to attach any interesting interpretation of the mean generation time.

A more reasonable definition is that generation time is defined conditional on that there are at least one secondary case. The question then arises how and if the generation time depends on the number of secondary cases. Is, e.g., the average time till the secondary cases the same regardless of the number of cases secondary to a primary case. Since generation times due to the same primary cases possibly are dependent there also remains the problem to derive the simultaneous distributions of the generation times of secondary cases related to the same primary case.

It turns out that it is a better alternative to define "backward" generation times and relate the term to the secondary cases rather than to the primary cases. In that case the generation time for any infected person is the time that has evolved since his infector was infected. This may seem like a simple rewording, but it is not and it turns out that it has essential implications.

3. A general formulation of an epidemic model

The first part of an epidemic model is to consider how many persons an infected person may infect and when the infections may occur. In order to do this we will consider two, possible random, entities, λ and K. Here λ is a non-negative (random) number that decides the "total amount of infectivity" spread by an infected person, K is a (random) positive measure, with total mass 1, defined on $[0, \infty[$ which measures how the infectiousness is distributed in time. The assumption that K is a random measure implies that it is not the same for each infected individual. It is chosen (independently for all persons) according to a distribution on all possible measures.

Observe that both λ and K are considered to be random and that they may be dependent.

A basic assumption is that for a given individual the number of possibly infectious contacts in the interval $I = [a, b[$ is Poisson distributed with mean $\lambda K(I)$ (conditional on λ and K). Let $K(t) = K([0, t[)$. For simplicity we assume that there exist a density so that

$$K(I) = \int_I \kappa(s)ds. \tag{3.1}$$

The full process of infectious contacts can be described by the following construction. Let χ be a homogeneous Poisson process with intensity 1. Then $\chi(I)$ is a random measure that counts the number of points in the interval I. We can define the new process that counts the number of infectious contacts in $I = [a, b[$ by

$$\xi(I) = \chi(]\lambda K(a), \lambda K(b)]). \tag{3.2}$$

An infectious contact results in a secondary case if the contact is taken with a susceptible person. In a simple homogeneous model the contacted persons are chosen randomly in the population. As explained above this part of the modelling will not be the concern of the present discussion, that focus on the inter-individual properties of the infection.

In this homogeneous model it is straight-forward that the basic reproduction number

$$R_0 = \mathrm{E}(\lambda). \tag{3.3}$$

This follows since K is assumed to have total mass 1.

4. Simultaneous distributions of generation times

Assume that a primary case has m secondary cases which occurs at times $(\tau_1, \ldots, \tau_m)$. It is well-known that conditional on $\chi([0, \lambda]) = m$, the time of the m events are distributed as $(\lambda U_1, \ldots, \lambda U_m)$ where $U_1, \ldots, U_m$ are independent uniformly distributed random variables.

This result can be translated to the times till secondary cases. The times $\tau_1, \ldots, \tau_m$ are distributed as

$$(K^{-1}(U_1), \ldots, K^{-1}(U_m)). \tag{4.1}$$

Here K^{-1} is the inverse of the function K. From this we can derive the simultaneous distribution of $\tau_1,\ldots,\tau_m$ (conditional on the number of secondary cases)

$$\Pr(\tau_1 \leq s_1, \ldots, \tau_m \leq s_m \mid m, K) = \prod_{j=1}^{m} K(s_j). \tag{4.2}$$

Given the actual random measure K the generation times are independent. If we loosen this restriction the generation times are no longer independent and

$$\Pr(\tau_1, \leq s_1, \ldots, \tau_m \leq s_m \mid m) = \int \prod_{j=1}^{m} K(s_j) dP(K). \tag{4.3}$$

Thus generation times generated by the same infector are exchangeable but not independent. However, they will all have the same mean (regardless of m).

$$\mathrm{E}(\tau \mid K) = \int_0^\infty a dK(a) = \int_0^\infty a k(a) da. \tag{4.4}$$

The overall mean, integrating over the distribution of K is

$$\mathrm{E}(\tau) = \int \left(\int_0^\infty a k(a) da \right) dP(K). \tag{4.5}$$

The covariance between two generation times related to the same primary case is

$$\mathrm{Cov}(\tau_1, \tau_2) = \mathrm{Var}(\mathrm{E}(\tau) \mid K) = \int \left(\int_0^\infty a k(a) da \right)^2 dP(K) - \left(\mathrm{E}(\tau)\right)^2. \tag{4.6}$$

Now we have to consider how to interpret the distribution derived above. It is the distribution (and mean) of the time till a secondary case (given that there exists at least one) if we choose a primary cases, which has at least one secondary case, with equal probability, ignoring how many secondary cases it has. We will denote the corresponding random variable by T_p and its distribution function by K_p. The density function of T_p is $k_p = K_p'$. It follows from the derivation above that

$$K_p(a) = \mathrm{E}(K(a)) = \int K(a) dP(K). \tag{4.7}$$

and

$$k_p(a) = \mathrm{E}(k(a)) = \int k(a) dP(K). \tag{4.8}$$

It is also possible, and perhaps more natural, to consider another generation time defined by choosing primary cases proportionally to the number of secondary cases, and then consider the distribution of the time till a secondary case. We denote the corresponding

random variable by T_s and its distribution function by K_s. The density function of T_s is $k_s = K_s'$.

The expected number of secondary cases will be proportional to λ. Thus

$$K_s(s) = \frac{\mathrm{E}_{\lambda K}(\lambda K(s))}{\mathrm{E}_\lambda(\lambda)}, \tag{4.9}$$

and the mean

$$\mathrm{E}(T_s) = \frac{\mathrm{E}_{\lambda K}(\lambda \int_0^\infty a k(a) da)}{\mathrm{E}_\lambda(\lambda)}. \tag{4.10}$$

The random variable T_s can be also be obtained by choosing a case at random and consider how long its infector has been infected.

Observe that the means of T_p and T_s are based on the same distributions, K. However, T_s is the result of a size-biased sampling procedure, where primary cases with many secondary cases are given a greater weight.

5. Some special models

5.1 *Constant infectiousness*

A common assumption in epidemic models is that each individual has a random latent time, X, a random infectious time Y, and a random infectivity γ. For simplicity we assume, as often is done, that these random variables are independent. Let g denote the density of the random variable X and h the density of Y. We will also use the distribution function of Y that is denoted by H.

We can rewrite these assumptions using the notation in the previous section:

$$\lambda = \gamma Y, \tag{5.1}$$

and

$$k(t) = \begin{cases} 1/\mathrm{Y} & \text{if } X < t \le X + Y, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

or with an alternative expression

$$k(t) = \frac{I(X < t \le X + Y)}{Y} \tag{5.3}$$

It follows directly that

$$R_0 = \mathrm{E}(\gamma)\mathrm{E}(Y). \tag{5.4}$$

Simple calculation yields that the mean of T_p is

$$\mathrm{E}(T_p) = \mathrm{E}(X) + \mathrm{E}(Y)/2. \tag{5.5}$$

The density of T_p equals $g * h_p$ (here $*$ stands for convolution) where

$$h_p(s) = \int_s^\infty \frac{h(a)}{a} da, \tag{5.6}$$

The density of T_s is $g * h_s$ where

$$h_s(s) = \frac{1 - H(s)}{\mathrm{E}(Y)}. \tag{5.7}$$

Thus

$$\mathrm{E}(T_s) = \mathrm{E}(X) + \frac{\mathrm{E}(Y^2)}{2\mathrm{E}(Y)}. \tag{5.8}$$

As an example we can consider the possibility that if Y is gamma-distributed with parameters $(\alpha, \alpha/\mathrm{E}(Y))$. Then

$$\mathrm{E}(T_s) = \mathrm{E}(X) + \mathrm{E}(Y)\frac{\alpha + 1}{2\alpha}. \tag{5.9}$$

It can be observed that if Y has an exponential distribution, i.e. $\alpha = 1$, then h_s is the density of an exponentially distributed random variable and

$$\mathrm{E}(T_s) = \mathrm{E}(X) + \mathrm{E}(Y). \tag{5.10}$$

5.2 *Time-varying infectiousness*

It many cases it is known that the infectiousness varies during the infectious period. However, it is often difficult to get a realistic model for this variation. To illustrate the possible effect of such variations we will here consider some simple cases. We will assume that the both the total infectiousness, i.e. λ and the random measure K depends on a real-valued random variable, γ. That is each individual has its own value of γ, and spreads the total infectivity λ_γ according to the measure K_γ.

In simple cases it is possible to find relations between T_p and T_s. Such a case occurs when both λ_γ and $K_\gamma(s)$, for all s, are increasing in γ. Then by applying Chebyshev's inequality (Hardy et al. (1934), pg 43) to the expressions (4.7) and (4.9), it follows for all s that

$$\frac{\mathrm{E}_\gamma(\lambda_\gamma K_\gamma(s))}{\mathrm{E}_\gamma(\lambda_\gamma)} \geq \mathrm{E}_\gamma(K_\gamma(s)). \tag{5.11}$$

This implies that the random variable T_s, in this case, is stochastically smaller than T_p. As a consequence of this $\mathrm{E}(T_s) \leq \mathrm{E}(T_p)$

To illustrate the connection between the total infectivity and how the spread is distributed over time we have chosen the following simple example. (Observe that in this example $K_\gamma(s)$ is not monotone in γ). Assume that

$$k_\gamma(a) = \gamma \exp(-\gamma a), \tag{5.12}$$

where γ has a $\Gamma(\alpha, \alpha/d)$-distribution, i.e., it has density function

$$\frac{\gamma^{\alpha-1}c^{\alpha}\exp(-c\gamma)}{\Gamma(\alpha)}, \tag{5.13}$$

where $c = \alpha/d$.

In order that the situation shall be interesting the total amount of infectivity has to correlate with the form of the infectious curve. Thus we assume that

$$\lambda_{\gamma} = \exp(-b\gamma). \tag{5.14}$$

Simple calculation now yields that

$$h_p(a) = d(1 + ad/\alpha)^{-(\alpha+1)}, \tag{5.15}$$

and

$$h_s(a) = d(1 + (a+b)d/\alpha)^{-(\alpha+1)}. \tag{5.16}$$

The corresponding means are

$$\mathrm{E}(T_p) = d\frac{\alpha}{\alpha-1}, \tag{5.17}$$

and

$$\mathrm{E}(T_s) = d\frac{\alpha}{\alpha-1} + b\frac{1}{\alpha-1}. \tag{5.18}$$

If $b < 0$ low total infectivity is associated with late secondary cases, i.e., $\mathrm{E}(T_s) < \mathrm{E}(T_p)$. If $b > 0$ the inequality is reversed.

6. Generation times observed from epidemics outbreaks

Statistical analysis of generation times that are derived from observations from an epidemic outbreak requires some care.

The first problem is that possible secondary cases may be lost due to immunity caused by previous infections. A consequence is that observed generation times will tend to be shorter in the intense stage of an epidemic. Let $S(t)$ be the number of susceptible individuals in the population at time t. A person that is infected at time z will then spread infectivity according to the function $\lambda S(z+a)K(a)$ at time a after z. If S decreases fast enough it may substantially reduce the possibilities for late infections and thus favor short generation times. If this is an effect that has to be considered or not depends on the time scale at which the epidemic spreads and the duration of the infectivity of an infected. In the start of an epidemic this effect may be disregarded.

A more serious problem is that the number of primary cases will grow at an exponential rate at the beginning. This means that at any time a large proportion of the infectors has not been able to spread all their infectivity. This implies that there is an over representation of early infected secondary cases. It is still possible to calculate a mean of the observed

ages of the cases that has occurred. We will here quote, without proof, a results from demography which there relates to mean maternal ages. It can be calculated as

$$\int ae^{-ra}k_s(a)da, \tag{6.1}$$

where r is the malthusian parameter that solves the equation

$$\int e^{-ra}k_s(a)da = 1, \tag{6.2}$$

(Jagers (1975) ch 8.4).

7. Serial times

As an alternative to generation times it has been suggested to consider times between observable events in primary cases and secondary cases. Figure 7.1 illustrates this. In the figure the time of infections as well as for some observable event, e.g., onset of symptoms of some kind are indicated. Let U denote the time between infection and the observable event, and let V denote the difference between the generation time and the observable event. Following the notation of the figure

$$T_g = U0 + V0. \tag{7.1}$$

However, the serial time is

$$S = V0 + U1. \tag{7.2}$$

It is clear that $\mathrm{E}(T_g) = \mathrm{E}(S)$ if $U0$ and $U1$ have the same distribution. However, in general, it is possible that T_g and S do not have the same distribution. It is natural to assume that $U0$ and $V0$ are dependent random variables since they are attached to the same individual and that $U0$ and $U1$ are independent since they come from two different individuals. If this is the case the distribution of T_g and S will not be the same. This is illustrated already by considering the variances:

$$\mathrm{Var}(T_g) = \mathrm{var}(S) + 2\mathrm{Cov}(U0, V0). \tag{7.3}$$

As a simple example we consider a three stage model for the inter-individual infectivity process. The duration of the latent period, X, is assumed to be exponential distributed with intensity λ_i. Then the infectiousness is constant during an exponentially distributed, with intensity λ_{inf}, infectious time. Symptoms are shown only after the end of the latent time with a delay, Y, that is exponentially distributed with intensity λ_d. The generation time for a secondary case is then the distribution of $X + Z$, where Z is exponentially distributed with intensity λ_{inf}. Simple calculations show that the distribution of the serial time is the distribution of $X + Z + Y_1 - Y$, where Y_1 is the delay for the secondary case. Thus $S = T_g + (Y_1 - Y)$. The added term, $Y_1 - Y$, has mean 0 and does not change the mean, but it has a real influence on the distribution of S.

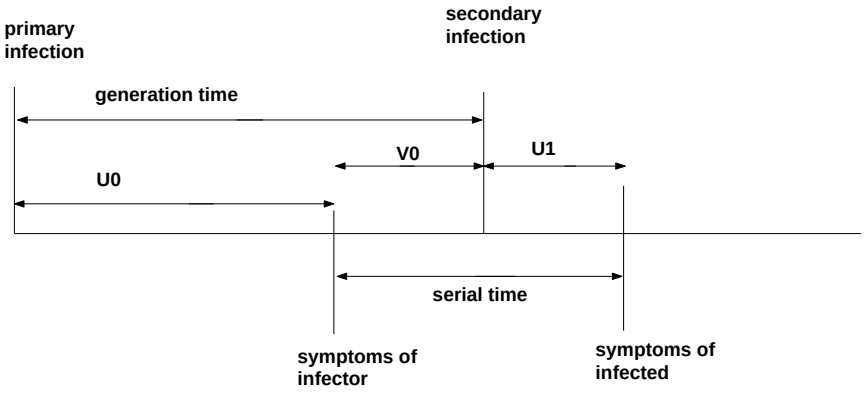

Figure 7.1. Generation time and serial time (observe that $V0$ might be negative)

Acknowledgements

This report has been financially supported by the DG Sanco project MODELREL and the DG Research project INFTRANS

References

Anderson, R., Fraser, C., Ghani, A. et al. (2004). Epidemiology, transmission dynamics and control of SARS: the 2002-2003 epidemic. *Phil. Trans. R. Soc. London. B* **359**, 1091–1105.

Anderson, R. and May, R. (1992). *Infectious Diseases of Humans*. Oxford University Press, Oxford.

Ferguson, N., Cummings, D., Cauchemez, S. et al. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214.

Fine, P. (2003). The Interval between Succesive Cases of an Infectious Disease. *Am J Epidemiol* **158**, 1039–1047.

Giesecke, J. (2002). *Modern Infectious Disease Epidemiology*. Arnold, London, second edition.

Haccou, P., Jagers, P. and Vatutin, V. (2005). *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press.

Hardy, G., Littlewood, J. and Pólya, G. (1934). *Inequalities*. Cambridge University Press, Cambridge.

Jagers, P. (1975). *Branching Processes with Biological Applications*. John Wiley & Sons.

Keyfitz, N. and Caswell, H. (2005). *Applied Mathematical Demography*. Statistics for Biology and Health. Springer, third edition.

Kimmel, M. and Axelrod, D. (2002). *Branching Processes in Biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer.

Wallinga, J. and Teunis, P. (2004). Different Epidemic Curves for Severe Acutre Respiratory Syndrome Reveal Similar Impacts of Control Measures. *Am J of Epidemiol* **160**, 509–516.