APPROXIMATING THE EPIDEMIC CURVE

A.D. BARBOUR
Univ. ZÜRICH

&

GESINE REINERT
OXFORD UNIV.

1. DETERMINISTIC EPIDEMIC MODELS.

Kermack & McKendrick, 1927.

N individuals, of whom $\boxed{S(t)}$ are susceptible at time t:

$S(t)$ treated as a continuous function of t:

$$-DS(t) = \frac{S(t)}{N} \int_0^\infty \beta(v)\{-DS'(t-v)\}\, dv, \qquad (1)$$

where $\boxed{\beta(v)}$ is the infectivity of an individual at time v after becoming infected, and $\boxed{R_0 := \int_0^\infty \beta(v)\, dv}$ is the mean number of individuals infected by a single infectious individual when all others are susceptible.

Writing $s(t) := S(t)/N$, (1) becomes

$$-Ds(t) = s(t) \int_0^\infty \beta(v) \{-Ds(t-v)\}\, dv,$$

which integrates to

$$-\log s(t) = \int_0^\infty \beta(v) \{1 - s(t-v)\}\, dv$$

if $s(-\infty) = 1$; or, equivalently,

$$s(t) = \exp\left\{-\int_0^\infty \beta(v) \{1 - s(t-v)\}\, dv\right\}. \qquad (2)$$

Letting $t \to \infty$ in (2) gives the 'final size' equation:

$$s(\infty) = \exp\left\{-\int_0^\infty \beta(v)\{1-s(\infty)\}\,dv\right\}$$

$$= e^{-R_0(1-s(\infty))}; \qquad (3)$$

the 'final size' is $N(1-s(\infty))$.

Interpretations of (3):

- 'force of infection'
- extinction probability of Poisson branching process.

2. THE POISSON BRANCHING PROCESS.

The Poisson branching process comes from looking backwards in time. Define, for K fixed,

$$X_i := I[\, i \text{ would infect}^* K, \text{ if } i \text{ became infected}\,], \quad i \neq K.$$

With homogeneous mixing, $\mathbb{P}[X_i = 1] = R_0/(N-1)$; and, if mixing is independent, so are the X_i. Hence the number of individuals that would potentially infect K is

$$\sum_{i \neq K} X_i \sim Bi(N-1, R_0/(N-1)) \approx Po(R_0).$$

Thus, at least for a moderate number of generations, the number of individuals born in a branching process with $Po(R_0)$ offspring distribution has distribution close to that of the number of individuals who, if infected, would lead to K being infected later.

- Forward and backward branching approximations
- Extinction probabilities and the final size.

3. A BRANCHING INTERPRETATION OF (2):

$$s(t) = \exp\left\{-\int_0^\infty \beta(v)\{1-s(t-v)\}\,dv\right\}.$$

Write $s(t) = \phi(ce^{\lambda t})$: (2) becomes

$$\phi(ce^{\lambda t}) = \exp\left\{-\int_0^\infty \beta(v)\{1-\phi(ce^{\lambda t}e^{-\lambda v})\}\,dv\right\},$$

or, writing $\theta = ce^{\lambda t}$,

$$\phi(\theta) = \exp\left\{-\int_0^\infty \beta(v)\{1-\phi(\theta e^{-\lambda v})\}\,dv\right\}. \quad (4)$$

If λ is chosen such that $\int_0^\infty e^{-\lambda v}\beta(v)\,dv = 1$, then

(4) is solved by $\phi(\theta) = E(e^{-\theta W})$, where

$W := \lim_{t\to\infty} e^{-\lambda t}Z(t)$ is a branching process limiting r.v.

Here, Z is a Crump-Mode-Jagers continuous time branching process, with offspring process the Poisson process with intensity $\beta(\cdot)$ on $\mathbb{R}_+$ and mean number of offspring $\int_0^\infty \beta(v)\,dv = R_0$, and λ is its Malthusian parameter.

- Interpretation: the backward branching process, including the time taken to infect K.

- Contrast with the forward branching process: λ and the intensity measures are the same, but the process distributions and extinction probabilities are not.

4. How is (2) justified probabilistically?

Example (not quite the same): the Reed–Frost discrete generation epidemic:

Graphical construction: start with a realization of the Bernoulli random graph $G(N, R_0/N)$. Choose an initially 'infected' vertex P at random. Then the d-neighbourhood of P with respect to graph distance is the set of individuals who have been infected by time d, and the final size of the epidemic is the size of the component containing P.

(a). Run the forward branching process until (if ever) about $\sqrt{N}$ individuals have been born.

Here, run for d generations, where d is such that $R_0^d = \lceil \sqrt{N} \rceil =: \psi_N \sqrt{N}$, with $1 \le \psi_N < R_0$. Label the individuals at random from $[N] := \{1, 2, \dots, N\}$. This (almost) gives the d-neighbourhood $\mathcal{N}_d(P)$. Write

$\mathcal{S}_m := [N] \setminus \mathcal{N}_m(P)$, the susceptibles at time m,

$\mathcal{I}_m := \mathcal{N}_m(P) \setminus \mathcal{N}_{m-1}(P)$, the infectives at time m:

write $S_m := |\mathcal{S}_m|$, $I_m := |\mathcal{I}_m|$.

* <u>If $I_d = 0$, the epidemic was small.</u> This event has probability close to ρ_F, the extinction probability of the forward branching process

(b) Run the backward branching process for a similar length of time.

Here, run for $d+n$ generations, $n \in \mathbb{Z}$. The backward process here is (almost) the same: just derived from $S_{d-1} \sim N$ individuals, instead of from exactly N. Label the individuals at random from S_{d-1}, calling the first K. Let $\bar{N}_{d+n}(K)$ denote the corresponding neighbourhood. Define the event

$$E_{d+n}(K) := \{ \bar{N}_{d+n}(K) \cap \mathcal{I}_d = \emptyset \}.$$

* On $E_{d+n}(K)$, $K \in S_{2d+n}$.

* $\mathbb{E}\left(\frac{S_{2d+n}}{N} \mid \mathcal{F}_d \right) = \left(1 - \frac{|N_{d-1}(P)|}{N}\right) \mathbb{P}[E_{d+n}(K) \mid \mathcal{F}_d] \sim \mathbb{P}[E_{d+n}(K) \mid \mathcal{F}_d]$.

(c) Use the Stein-Chen method to approximate $\mathbb{P}[E_{d+n}(k) \mid \mathcal{F}_d]$.

Because of random labelling,

$$\mathbb{P}[E_{d+n}(k) \mid \mathcal{F}_d] \sim \mathbb{E}\{e^{-|\bar{N}_{d+n}(k)| \cdot |\mathcal{I}_d|/N} \mid \mathcal{F}_d\}.$$

By branching asymptotics, as $N \to \infty$,

$$|\bar{N}_{d+n}(k)| \sim \bar{W} R_0^{d+n} = \bar{W} \psi_N \sqrt{N} R_0^n ;$$

$$|\mathcal{I}_d| =: W(d) R_0^d = W(d) \psi_N \sqrt{N}, \text{ with } W(d) \sim W \text{ as } N \to \infty;$$

hence

$$\mathbb{E}\left(\frac{\xi_{k,d+n}}{N} \mid \mathcal{F}_d\right) \sim \mathbb{P}[E_{d+n}(k) \mid \mathcal{F}_d] \sim \mathbb{E}\{e^{-\bar{W} W(d) \psi_N^2 R_0^n} \mid \mathcal{F}_d\}$$

$$= \phi(c_N e^{\lambda n}), \text{ with } c_N = \psi_N^2 W(d), \ \lambda = \log R_0 \text{ and } \phi(\theta) = \mathbb{E}(e^{-\theta \bar{W}}).$$

(d) Now observe that $\mathbb{E}\{(\frac{S_{2dtn}}{N})^2 \mid \mathcal{G}_d\} \sim \mathbb{P}[E_{dtn}(k) \cap E_{dtn}(k') \mid \mathcal{G}_d]$, for two independently chosen random individuals k, k'.

The corresponding neighbourhoods $\bar{N}_{dtn}(k)$ and $\bar{N}_{dtn}(k')$ are again realized by independent labelling, and it is easy to conclude that

$$\mathbb{P}[E_{dtn}(k) \cap E_{dtn}(k') \mid \mathcal{G}_d] \sim (\mathbb{P}[E_{dtn}(k) \mid \mathcal{G}_d])^2.$$

Hence $\mathrm{var}(N^{-1} S_{2dtn} \mid \mathcal{G}_d) = o(1)$ as $N \to \infty$, and so, conditional on $\mathcal{G}_d$,

$$\left| N^{-1} S_{2dtn} - \phi(c_N e^{\gamma n}) \right| \longrightarrow_1 0 \text{ as } N \to \infty.$$

NB. $c_N = \psi_N^2 W(d)$; $\psi_N \in [1, R_0]$ is non-random; $\mathbb{P}[W(d) > 0] \sim 1 - \varrho_F$.

For the Kermack-McKendrick theorem,

$\mathcal{I}_m$ is replaced by $\tilde{\mathcal{I}}_\tau$, the unborn children of those born before τ, the time at which $\lceil \sqrt{N} \rceil$ individuals have been born, and their birth times, b_i;

$\mathcal{N}_n(K)$ is replaced by $\overline{\mathcal{B}}_u(K)$, the history of the backward branching process up to time u, including all birth times, $\bar{b}_j$;

A common label for $i \in \tilde{\mathcal{I}}_\tau$ and $j \in \overline{\mathcal{B}}_u(K)$ only implies infection before time $\tau + u$ if $b_i - \tau \le u - \bar{b}_j$

Stein-Chen still handles this.

empirical distns. covered by branching asymptotics.