# SAMSI Ecology Class

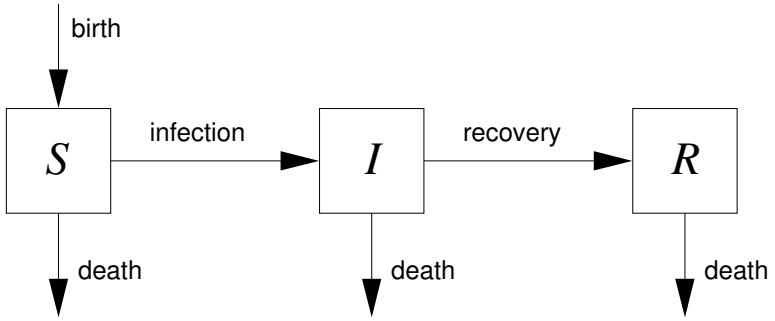
## Modeling the Spread of Infectious Diseases: II. Stochastic Theory

Alun Lloyd

# Plan for Today's Lecture

- Very brief recap of last week
- Data: Motivating us to look at stochastic models
- Results from simulations of stochastic SIR models
- Invasion of infection: branching process theory
- Persistence of endemic infection:
  - estimating variability about endemic equilibrium
  - moment equations (and moment closure) as one approach
- Kurtz's results

# Recap: Deterministic SIR models




$$\frac{dS}{dt} = \mu N - \frac{\beta SI}{N} - \mu S$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Can consider SIR with or without demography (epidemic vs endemic settings)

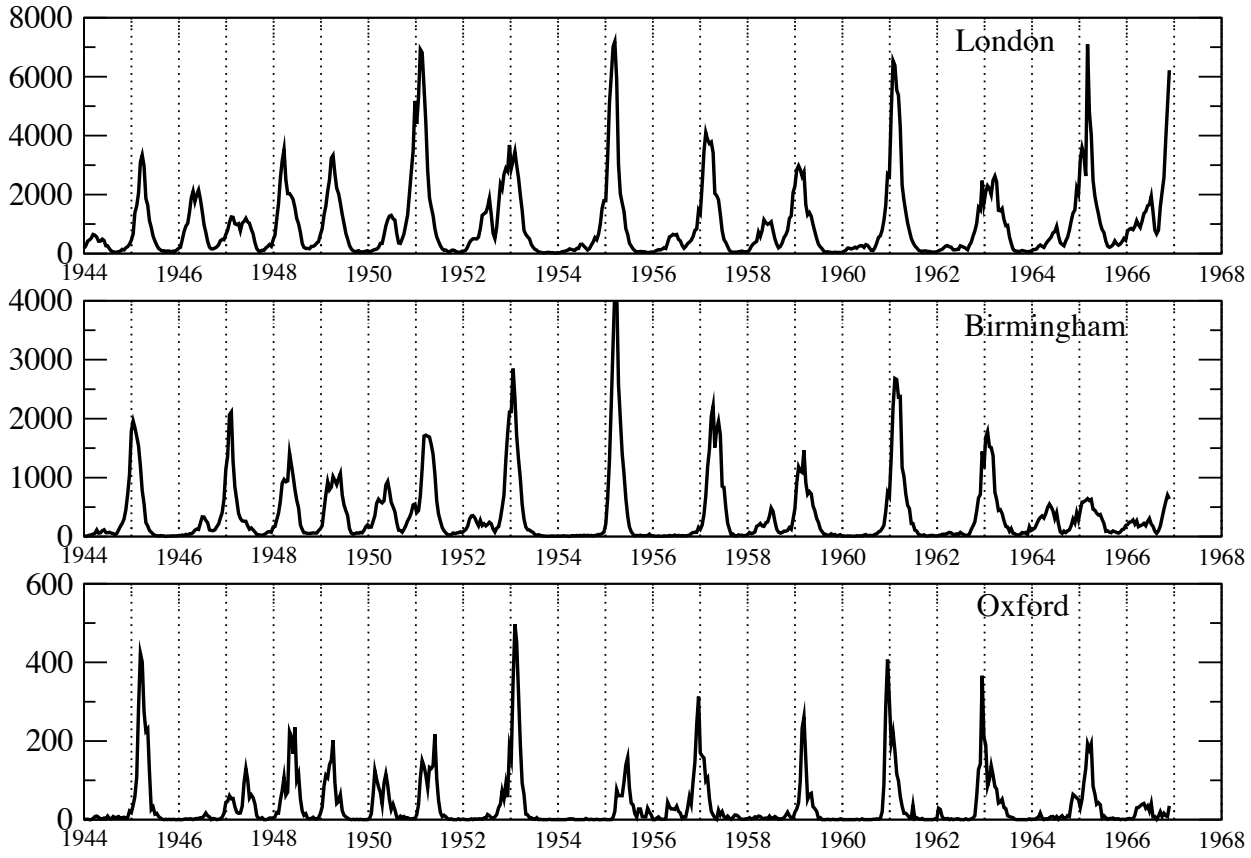
- Invasion behavior understood by assuming $S \approx N$ , which converts the nonlinear transmission term $\beta SI/N$ into the linear $\beta I$ **linear approximation**

  Leads to the notion of basic reproductive number $R_0$, and the invasion threshold $R_0 = 1$

- Endemic behavior: existence of a stable endemic equilibrium if $R_0 > 1$

# Data: Measles Incidence vs SIR model with Demography?

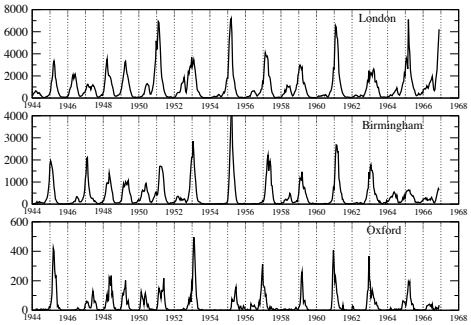
Measles incidence in 3 British cities before the introduction of measles vaccine (late 60s)



# Data: Measles Incidence vs SIR model with Demography?

In the data, we see:

- **Sustained oscillations.** Measles epidemics occur with a two-yearly (biennial) period in most cases. In some instances, annual epidemics are seen, while in others, epidemics occur every 3 years (triennial).
- **Some Irregularities.** The peak incidences differ between outbreaks in the same city.
- **Breaks in the chain of infection.** The infection can undergo 'fadeout' following an epidemic as the number of infectives falls to low levels in the 'troughs' between epidemics.




SIR model (or SEIR model) with demography instead predicts damped oscillations to equilibrium

What are we missing?

1. Stochastic effects
2. Seasonal forcing : transmission is higher during school terms than vacations

# Stochasticity in Disease Models

Individuals are discrete

Rather than having continuous flow (e.g.) from S to I, have discrete transitions

Reinterpret the transition rates of ODE model (state space $\mathbb{R}^2$ ) as transition rates of a continuous-time Markov process (state space $\{0, 1, 2, \dots\}^2$ )

| Event | Transition | Rate at which event occurs | Probability of transition in time interval $[t, t+dt]$ |
|---|---|---|---|
| Birth | $S \to S+1$ | $\mu N$ | $\mu N dt$ |
| Susceptible death | $S \to S-1$ | $\mu S$ | $\mu S dt$ |
| Infection | $S \to S-1, \quad I \to I+1$ | $\beta SI/N$ | $(\beta SI/N)dt$ |
| Recovery | $I \to I-1$ | $\gamma I$ | $\gamma I dt$ |
| Infectious death | $I \to I-1$ | $\mu I$ | $\mu I dt$ |

Stochastic effects due to finite population size: demographic stochasticity

"intrinsic stochasticity"

# Stochasticity in Disease Models: Simulation Results

We have seen how to simulate this sort of model (Gillespie algorithm)

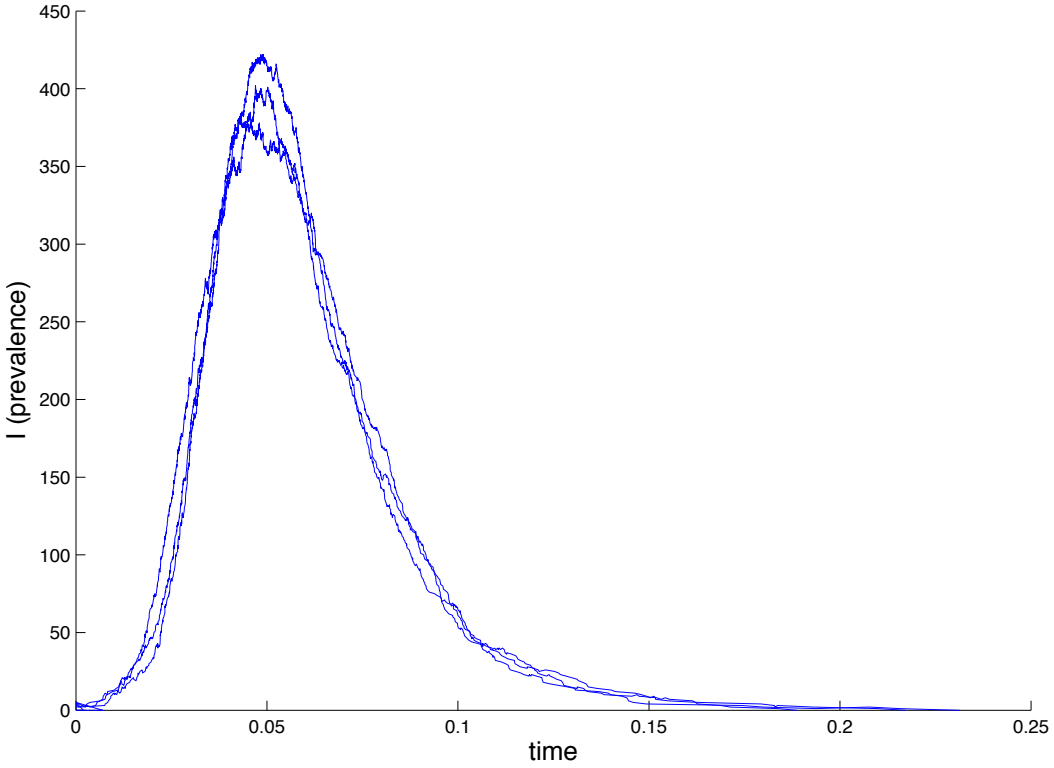
Five realizations of stochastic SIR model without demography:

$N$=1000, $\beta$=200, $\gamma$=100

One initial infective, rest of population susceptible

Rapid extinction seen in 2 of the 5 realizations

The other 3 realizations follow trajectories that are much like the ODE model, but exhibit some variability

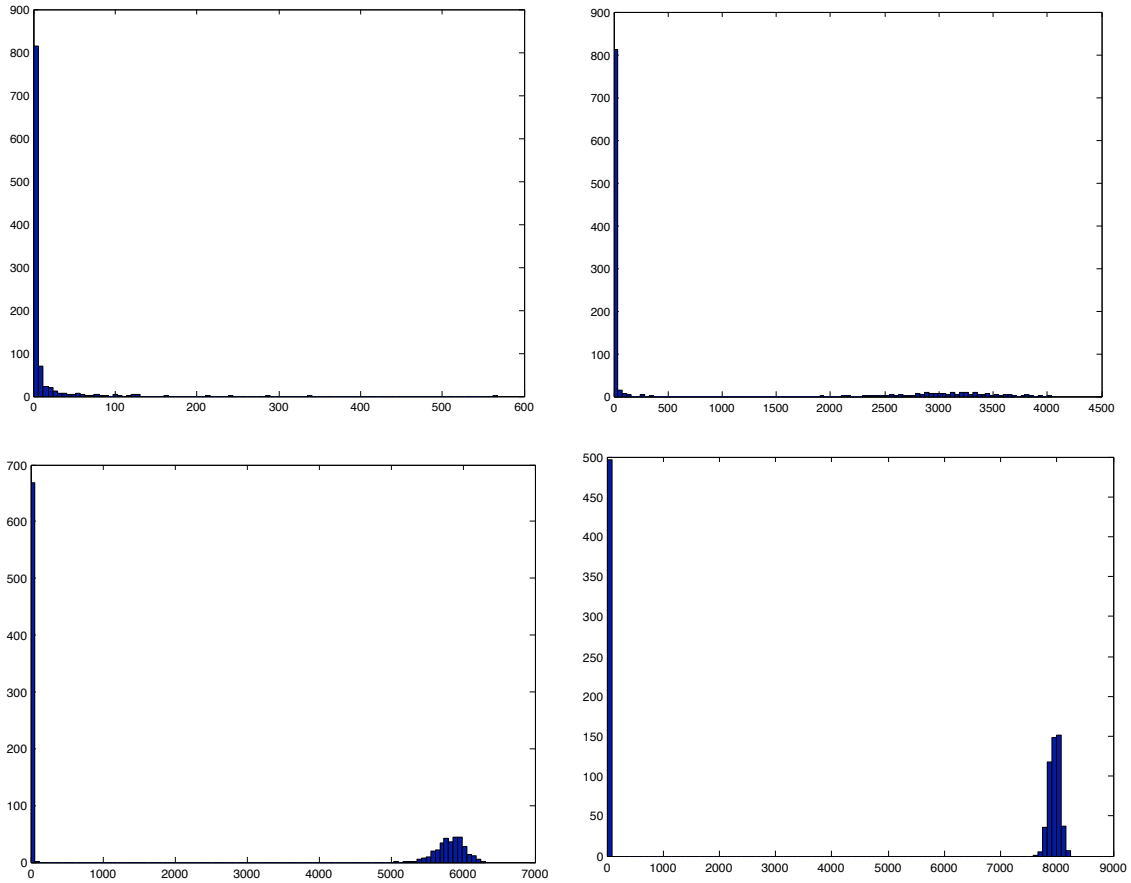


# Outbreak Size Distributions

SIR model without demography, $N$=10000, one initial infective

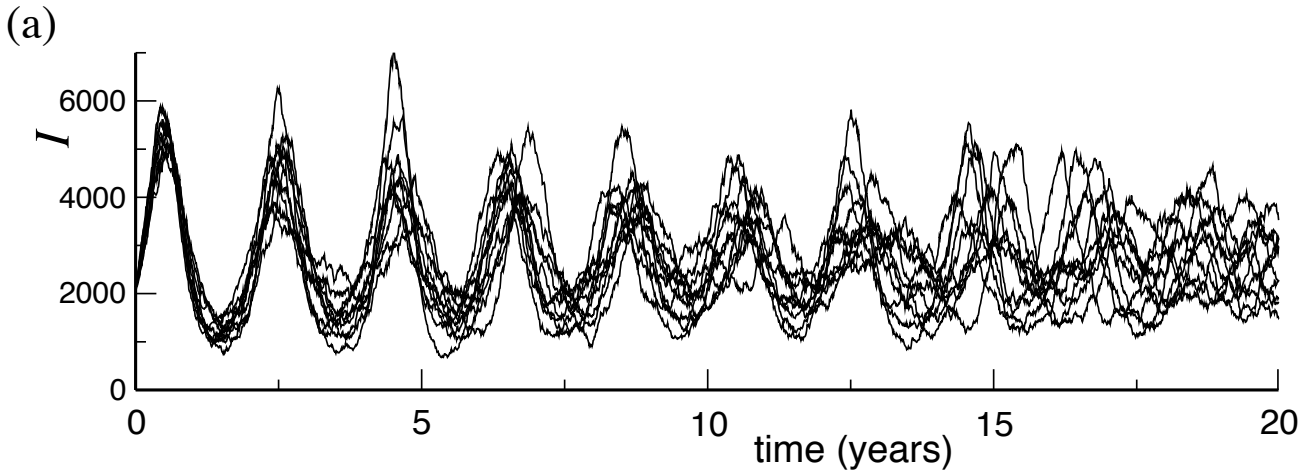
Outbreak size distributions across 1000 realizations for $R_0$= 0.9 , 1.2, 1.5 and 2.0

Note large peak at size 1
(no secondary cases!)



# Stochasticity in Disease Models: Simulation Results

Ten realizations of stochastic SIR model with demography:




$N = 10^7$, parameter values $\mu = 1/70 \text{ year}^{-1}$, $\gamma = 50 \text{ year}^{-1}$, $\beta = 750 \text{ year}^{-1}$

Initial conditions chosen to be reasonably close to endemic equilibrium

Rather than settle down into endemic equilibrium, individual realizations continue to fluctuate due to stochasticity

fluctuations are relatively large, despite total population of ten million!

# When is Demographic Stochasticity Important?

Simulation suggests demographic stochasticity can be important during an initial invasion phase and about endemic equilibrium

- Can we understand when demographic stochasticity will be important?
- Can we estimate impact of demographic stochasticity?
- What is the relationship between stochastic and deterministic model?

(When do we have to use a stochastic description?)

# Key Point: Interruption in Transmission Chain

If we have few infectives, there would be a chance that they all recover before transmitting infection

e.g. one infective, giving rise to infections at per-capita rate $\beta S/N$, but recovering at per-capita rate $\gamma$

What is the probability that they recover before transmitting infection?

(competing hazards calculation)

# When is Demographic Stochasticity Important?

Transmission chain can be broken even if $R_0 > 1$

In a stochastic model, invasion and persistence is no longer guaranteed by $R_0 > 1$

Interruption of transmission most likely when number of individuals is "small"

- During the initial invasion period
- Following a large outbreak (e.g. initial invasion outbreak) when *I* crashes
- Endemic phase

We shall first look at initial invasion period, using Galton-Watson branching process theory

# **Probability Generating Function** (pgf)

Let *X* be a random variable on the non-negative integers with p.m.f. $p_i$

Probability generating function $G_X(s) := \mathrm{E}(s^X) = \sum_{i=0}^{\infty} s^i \, p_i$

We shall drop the subscript *X* if there is no possibility of confusion

Pgf is a way of summarizing an entire probability distribution using a single function of a "dummy variable" *s*.

Example: Poisson random variable *X* with mean $\lambda$

$$p_i = \frac{\lambda^i e^{-\lambda}}{i\,!}$$

$$G_X(s) = \sum_{i=0}^{\infty} s^i \, \frac{\lambda^i e^{-\lambda}}{i\,!}$$

$$= e^{-\lambda} \sum_{i=0}^{\infty} \frac{(s\lambda)^i}{i\,!}$$

$$= e^{-\lambda} e^{s\lambda} = e^{\lambda(s-1)}$$

# Properties of Probability Generating Functions

1. $G_X(0) = \text{Prob}(X = 0)$
2. $G_X(1) = \sum_{i=0}^{\infty} p_i = 1$
3. $G_X(s)$ is concave up on [0,1]
4. Derivative with respect to $s$ : $G'_X(1) = \left( \sum_{i=0}^{\infty} i\, s^{i-1}\, p_i \right) \Big|_{s=1} = \sum_{i=0}^{\infty} i\, p_i = \mathrm{E}(X)$
5. Variance: $G''_X(1) + G'_X(1) - \{G'_X(1)\}^2 = \mathrm{Var}(X)$
6. If $X$ and $Y$ are independent random variables, then $G_{X+Y}(s) = G_X(s) G_Y(s)$
   result generalizes to sum of $n$ independent r.v.
7. Let $X_1, \ldots X_N$ be i.i.d. r.v. with common generating function $G_X(s)$, and let $N$ be a non-negative random variable that is independent of the $X_i$ with pgf $G_N(s)$ then $Y = X_1 + \ldots + X_N$ has generating function $G_Y(s) = G_N\left(G_X(s)\right)$

Result 6 describes the pgf of a sum of a fixed number of iid random variables, while result 7 describes the pgf for the sum of a random number of iid random variables

# Examples

Geometric r.v. thinks about successive independent Bernoulli trials, each of which has probability *p* of success, probability 1-*p* of failure, and timing of first success

Two slightly different versions:

*X* : number of failures before first success

*Y* : number of trial on which first success happens

*Y* = *X* + 1

$\mathrm{Prob}(X = i) = (1-p)^i p$ leads to $G_X(s) = \dfrac{p}{1 - s(1-p)}$

$\mathrm{Prob}(Y = i) = (1-p)^{i-1} p$ gives $G_Y(s) = \dfrac{ps}{1 - s(1-p)}$

Note: could have derived pdf of *Y* from that of *X*:

Y = *X* + 1 = *X* + *Z*, where *Z* is a r.v. that equals one. Pgf of *Z* is *s*, so

$G_X(s) = s\, G_Y(s)$

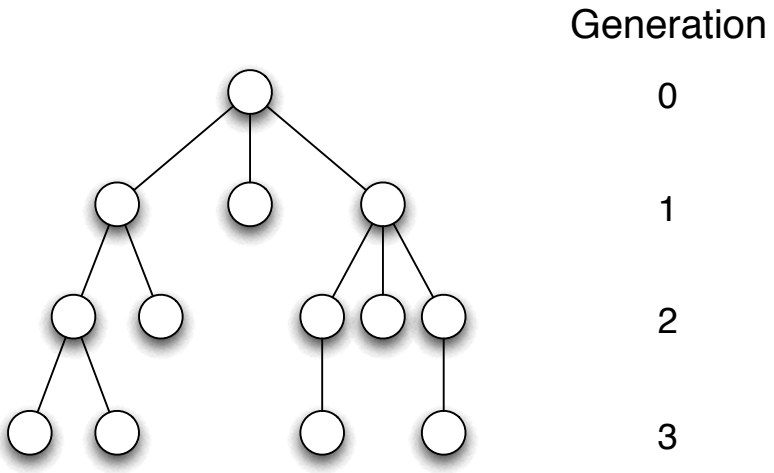
# Galton-Watson Branching Process

Consider chain of infection that starts with one initial infective

Ask the question: does this chain of infection go extinct or not?

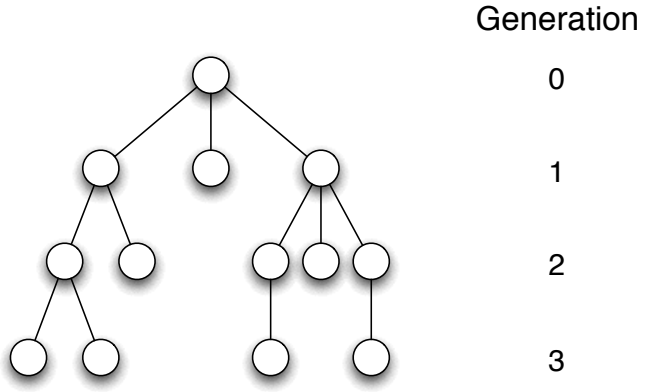
Employ a discrete time description in terms of "generations"

loses temporal information, but this is fine if we are interested in whether infection spreads or goes extinct




Anticipate (on average) a geometric growth/decay process, depending on $R_0$ being greater/less than 1; Number of individuals in generation $n$ to be $R_0^n$

# Galton-Watson Branching Process




Assume numbers of "offspring" (secondary infections), $Z$ , of individuals are independent, identically distributed random variables

Important comment: this assumption means we neglect depletion of susceptibles (which would reduce number of secondary infections as epidemic proceeds)

*Branching process is describing the early stages of an epidemic*
*as in the deterministic analysis of invasion, we use a linear approximation*

Offspring distribution: Common probability mass function, $p_i$ , and pdf $G_Z(s)$

We assume $p_0 > 0$ and that no $p_i = 1$ (removes trivial cases)

$X_n$ : total number of offspring in generation $n$

Key questions: what happens to $X_n$ as $n \to \infty$ ? What is the probability of lineage going extinct?

# Galton-Watson Branching Process: theory

Let $Z_{n,i}$ be r.v. describing offspring in generation *n*
for *i* 'th individual in generation *n*-1

Number of individuals in generation *n* is given by $X_n = \sum_{i=1}^{X_{n-1}} Z_{n,i}$

Because offspring are iid, we have the sum of $X_{n-1}$ iid rv [random sum]

So: $G_{X_n}(s) = G_{X_{n-1}}(G_Z(s))$

Apply this result repeatedly, get $G_{X_n}(s) = G_Z(G_Z(\cdots(G_Z(s))\cdots))$

$= G_Z^{(n)}(s)$

Generating function for total number of individuals in generation *n* is the *n*-fold iterate of the pgf of the offspring distribution

# **Extinction Probability for G-W Branching Process** (sketch)

Define $s_n$ = Prob( $X_n = 0$ )

Easy to see that $s_0 \leq s_1 \leq s_2 \cdots \leq 1$ : non-decreasing sequence, bounded above by 1

Therefore $\lim_{n\to\infty} s_n = s_\infty$ exists and is ≤ 1

Condition on number of individuals in generation 1:

$$
\begin{aligned}
s_n &= \sum_{k=0}^{\infty} \mathrm{Prob}(X_n = 0 | X_1 = k)\mathrm{Prob}(X_1 = k) \\
&= \sum_{k=0}^{\infty} s_{n-1}^k \, p_k \\
&= G_Z(s_{n-1})
\end{aligned}
$$

Probability that *k* lineages are all extinct at generation *n* – 1 , each starting from 1 individual

Limiting value $s_\infty$ satisfies the fixed point equation $s_\infty = G_Z(s_\infty)$

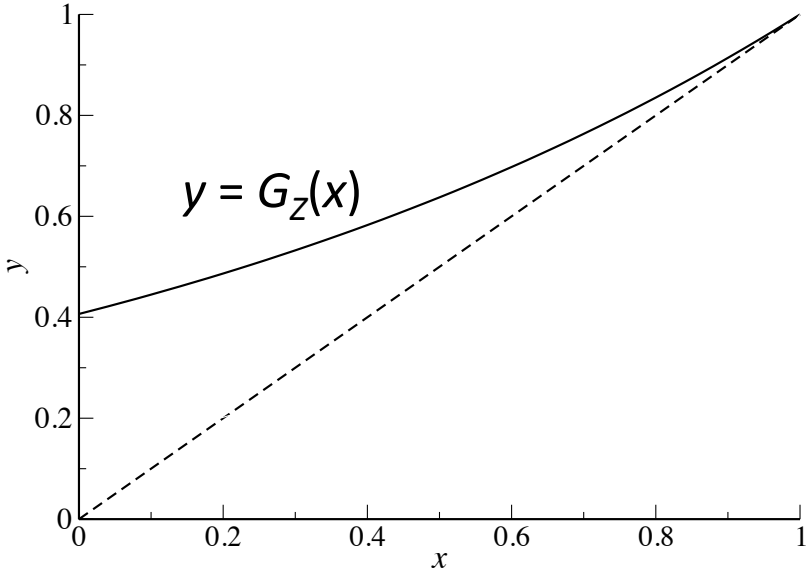
# **Extinction Probability for G-W Branching Process** (sketch)

Solve $G_Z(s) = s$ : look for intersections between graphs of $y = x$ and $y = G_Z(x)$

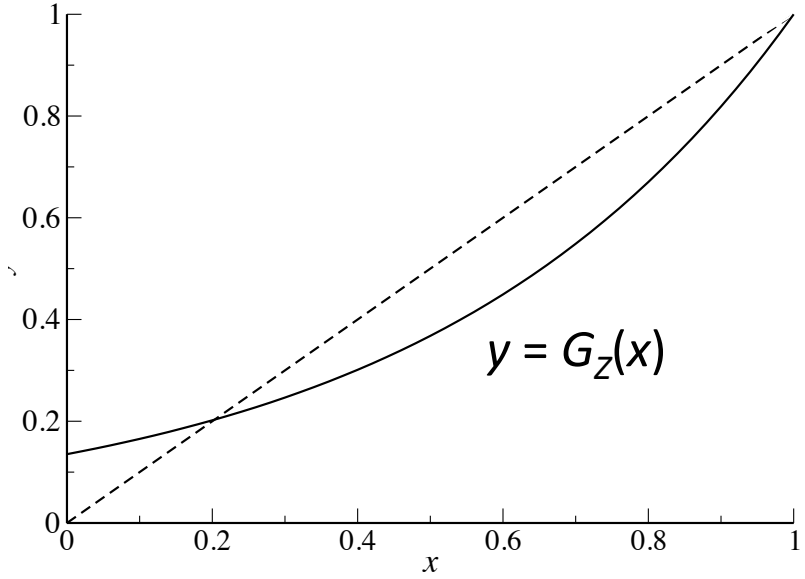
Recalling: $G_Z(0) = p_0 > 0$ , $G_Z(1) = 1$ , $G'_Z(1) = \mathrm{E}(Z) = R_0$ , $G_Z(x)$ is concave up …

… get two qualitatively different pictures depending on whether $R_0 < 1$ or $> 1$ :

$R_0 < 1$




$R_0 > 1$



# Key Result for G-W Branching Process

Probability of eventual extinction of the process is given by the smallest non-negative root of $G_Z(s) = s$ ,

This probability is:

- 1 if average number of offspring (i.e. $R_0$) is less than one
- less than 1 if $R_0$ is greater than one

# Offspring Distribution for SIR Model?

For a given infective, secondary infections are produced at rate $\beta S/N \approx \beta$ (linear approx.)
(Poisson process with rate $\beta$)

Duration of infection is exponentially distributed with mean $1/\gamma$

If duration of infection was fixed at $1/\gamma$, number of offspring would be Poisson distributed with mean $\beta/\gamma = R_0$

Calculate offspring distribution by conditioning on duration of infection:

$$\text{Prob}(Z = k) = \int_0^\infty \text{Prob}(Z = k|T = \tau)\, f_T(\tau)\, d\tau$$

where $f_T(t)$ is the pdf of the distribution of durations of infection

$$\text{Prob}(Z = k) = \int_0^\infty \frac{(\beta\tau)^k e^{-\beta\tau}}{k!}\, \gamma e^{-\gamma\tau} d\tau$$

Reduces to an integral of $\tau^k \exp(-(\beta+\gamma)\tau)$, but perhaps easier to directly calculate pgf:

$$G_z(s) = \sum_{k=0}^{\infty} s^k \int_0^\infty \frac{(\beta\tau)^k e^{-\beta\tau}}{k!}\, \gamma e^{-\gamma\tau} d\tau$$

# Offspring Distribution for SIR Model?

$$G_z(s) = \sum_{k=0}^{\infty} s^k \int_0^{\infty} \frac{(\beta\tau)^k e^{-\beta\tau}}{k!} \gamma e^{-\gamma\tau} d\tau$$

Switch order of summation and integration and then do summation (exponential series)

$$G_z(s) = \int_0^{\infty} e^{s\beta\tau} e^{-\beta\tau} \gamma e^{-\gamma\tau} d\tau$$

Gives

$$\begin{aligned} G_z(s) &= \frac{\gamma}{\gamma - \beta(s-1)} \\ &= \frac{1}{1 - R_0(s-1)} \end{aligned}$$

which we recognize as the pgf of a geometric distribution with mean $R_0$

# Application of G-W bp Result to SIR Model

Probability of eventual extinction of the process is given by the smallest non-negative root of $G_Z(s) = s$ ,

This probability is:

- 1 if average number of offspring (i.e. $R_0$) is less than one
- less than 1 if $R_0$ is greater than one

For our SIR model, the number of secondary infections is geometrically distributed, with mean $R_0$

$$G_Z(s) = \frac{1}{1 + R_0(1 - s)}$$

Solving $G_Z(s) = s$ gives extinction probability as $s = 1$ or $s = 1/R_0$

If $R_0 < 1$, $s = 1$ is relevant solution; if $R_0 > 1$ it's $s = 1/R_0$

# Linear vs Nonlinear Model

Summary for **linear** model with geometric offspring distribution:

Starting with one infective, then

if $R_0 < 1$ , epidemic process dies out with probability 1

if $R_0 > 1$ , epidemic process dies out with probability $1/R_0$

In the real world, the linear approximation breaks down once there is appreciable depletion of susceptibles…

But if this has happened, the number of infectives will have grown sufficiently that will be in a regime where the dynamics is reasonably well described by the deterministic model plus some noise

In the full model, if early extinction occurs (i.e. during the linear phase), we say there has been a “minor outbreak”. If we avoid early extinction, we say there has been a “major outbreak”

# More Thoughts on Minor/Major Outbreaks

What are the probabilities of minor and major outbreaks if we start with some general initial number of infectives? (Still assumed to be relatively small.)

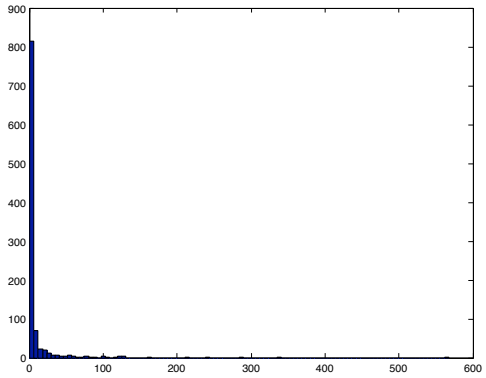
How do the components of the outbreak size distribution representing minor and major outbreaks change as the population size $N$ is increased?

# Outbreak Size Distributions, revisited

Returning to our earlier pictures of outbreak size distributions

$R_0 = 0.9$

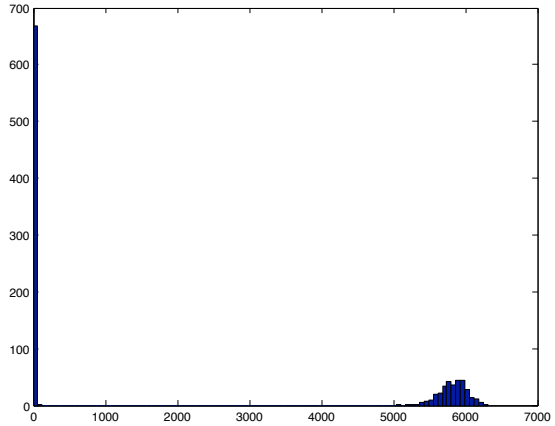
Only have minor outbreaks
outbreak size distribution is
J-shaped

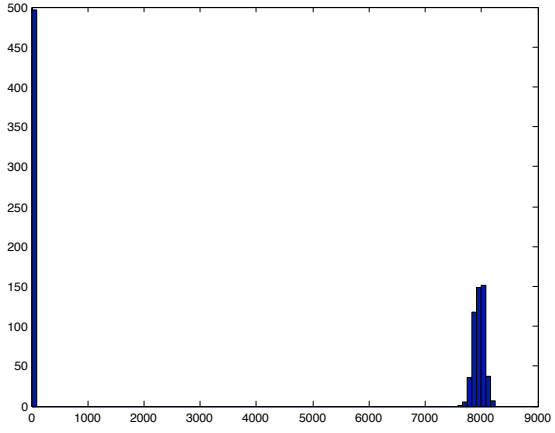



$R_0 > 1$

Get a mix of minor and major outbreaks

Major outbreak distribution is centered on value predicted by deterministic analysis, and variance decreases with $R_0$



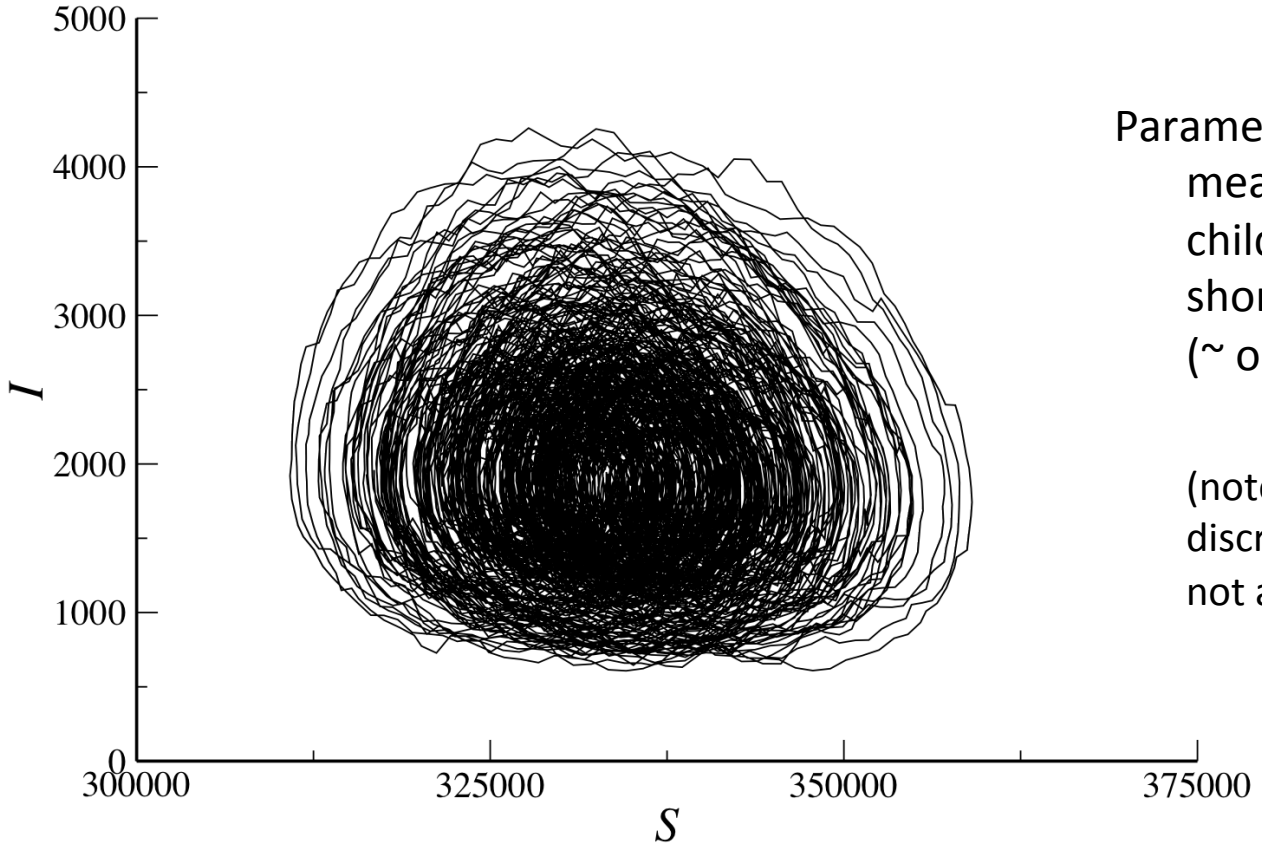



# Endemic Dynamics of Stochastic SIR Model

Deterministic model approached endemic equilibrium when $R_0>1$

Stochastic model: trajectories fluctuate around deterministic equilibrium

Number of infectives appears to stays well away from zero if $N$ is large enough (here $N$ = 5 million)




Parameters representative of measles: highly infectious childhood disease with a short infectious period (~ one week)

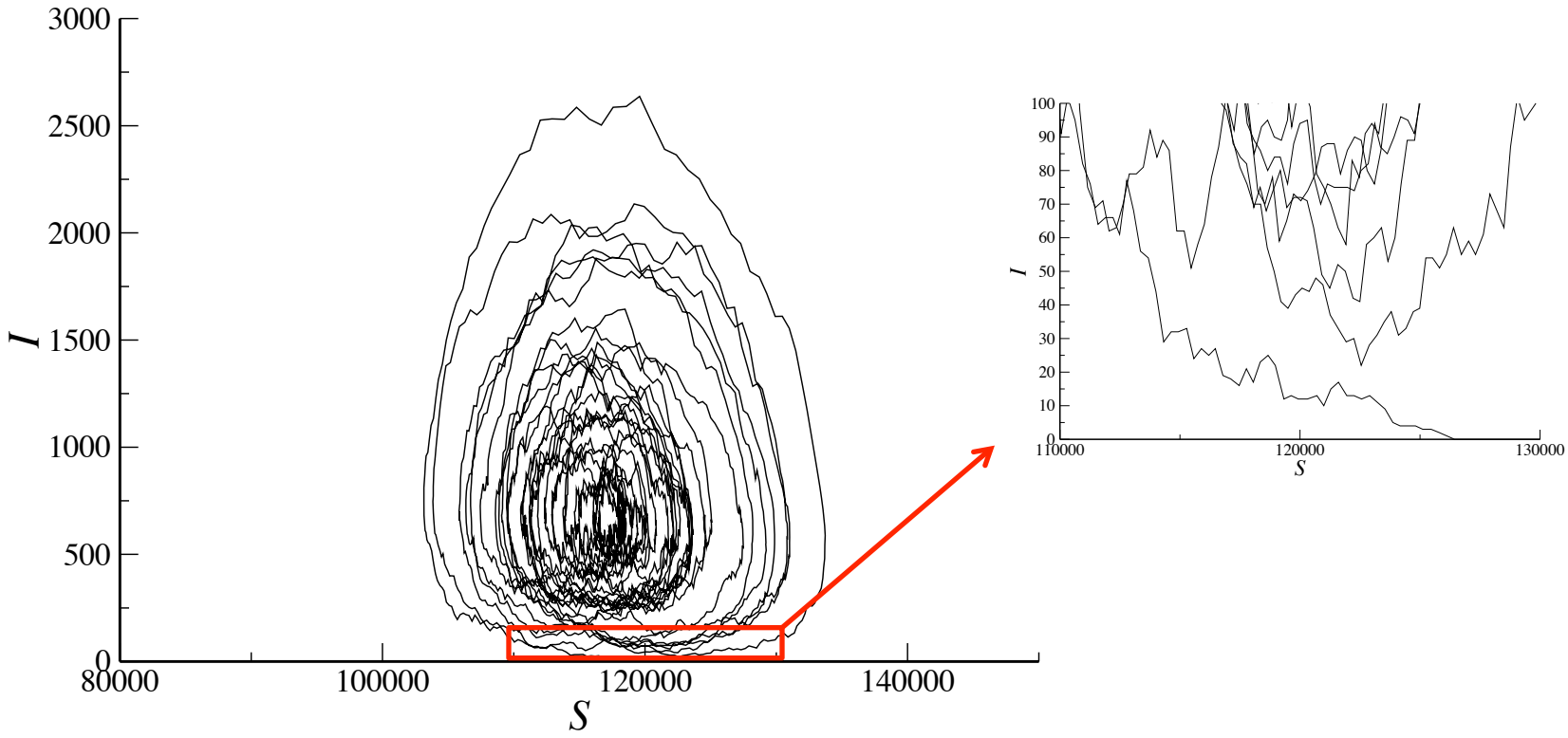
(note: trajectories are shown at discrete observation times, so not all jumps are shown)

# Dynamics of Stochastic SIR Model, II

Qualitatively different picture when *N* is 1.75 million:

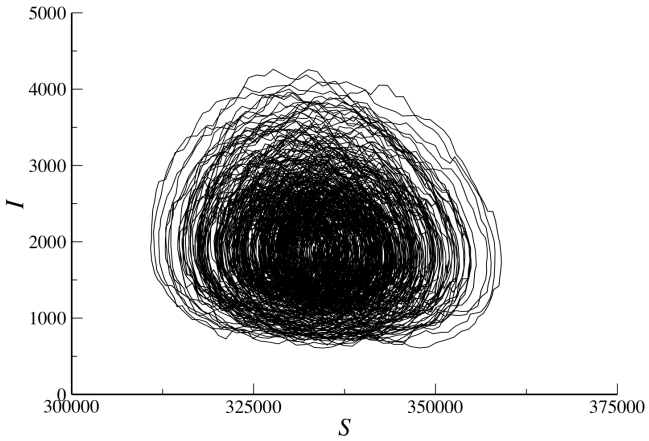
Oscillations bring *I* close to 0, with trajectory hitting horizontal axis

Extinction of infection due to stochastic effects: fadeout



# Dynamics of Stochastic SIR Model, III

Back to $N$ = 5 million




Number of infectives stays well away from zero?

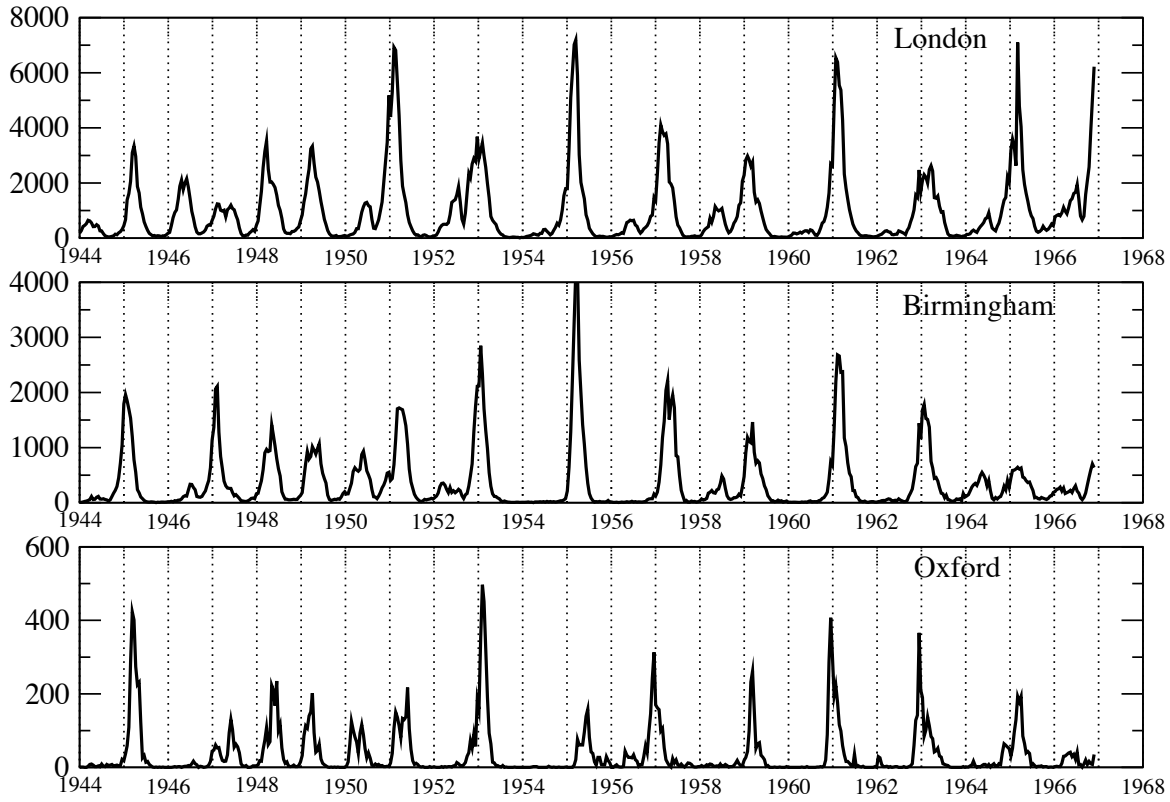
Not exactly: infection goes extinct with probability one for any $N$, although timescale may be astronomical (scales like exp(-a$N$) )

Have a quasi-stationary distribution about the endemic equilibrium, but eventually leave…

After extinction, trajectories approach the unstable equilibrium of the deterministic system

# Measles: Population Size Dependent Dynamics

Fortnightly cases of measles in three British cities: London (3-5 million people), Birmingham (1 million), Oxford (100 000), in the era before mass-vaccination




Number of cases in Oxford often falls to zero between outbreaks

Fadeouts seen in data

# Measles: Critical Community Size

Fadeouts rarely occurred for cities of more than a few hundred thousand inhabitants (Bartlett, 1950s)

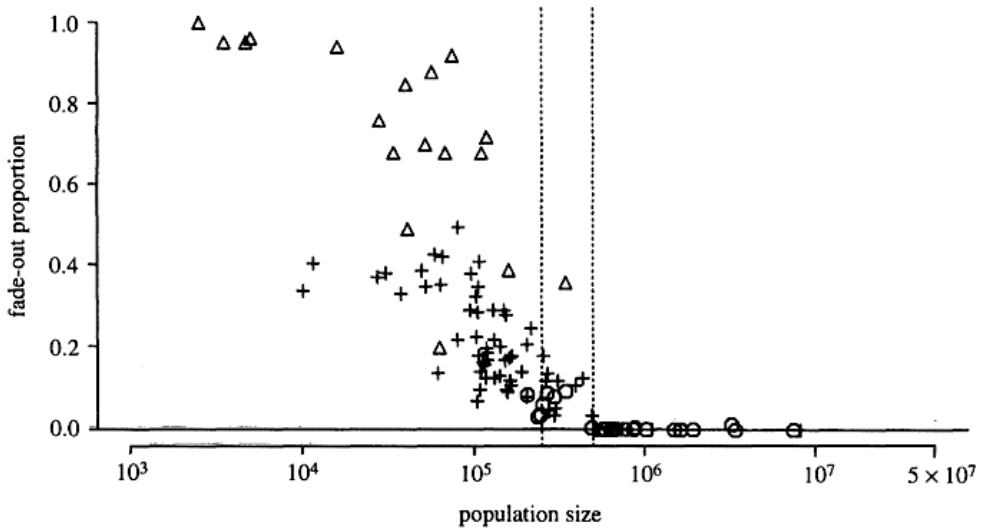



Figure 1. Empirical estimates of critical community size. Horizontal axis shows population size; vertical axis shows 'fade-out proportion', the fraction of months in each sample with no reported cases of measles. The vertical dotted lines represent the empirical estimate of critical community size range (Bartlett 1957, 1960*a*), 250000–500000. Data from Black (1966) (△, islands) Bartlett (1960*a*) (○, U.S. and Canadian cities), and Shaw (1990) (+, British cities; data originally from OPCS (1948–68)). Although there are many possible sources of variation in the data, including different reporting rates, *per capita* birth rates and levels of outside epidemiological contact, all the data fall roughly along the same curve.

Bolker & Grenfell, (1995) Phil. Trans. R. Soc. Lond. B 348, 309-320.

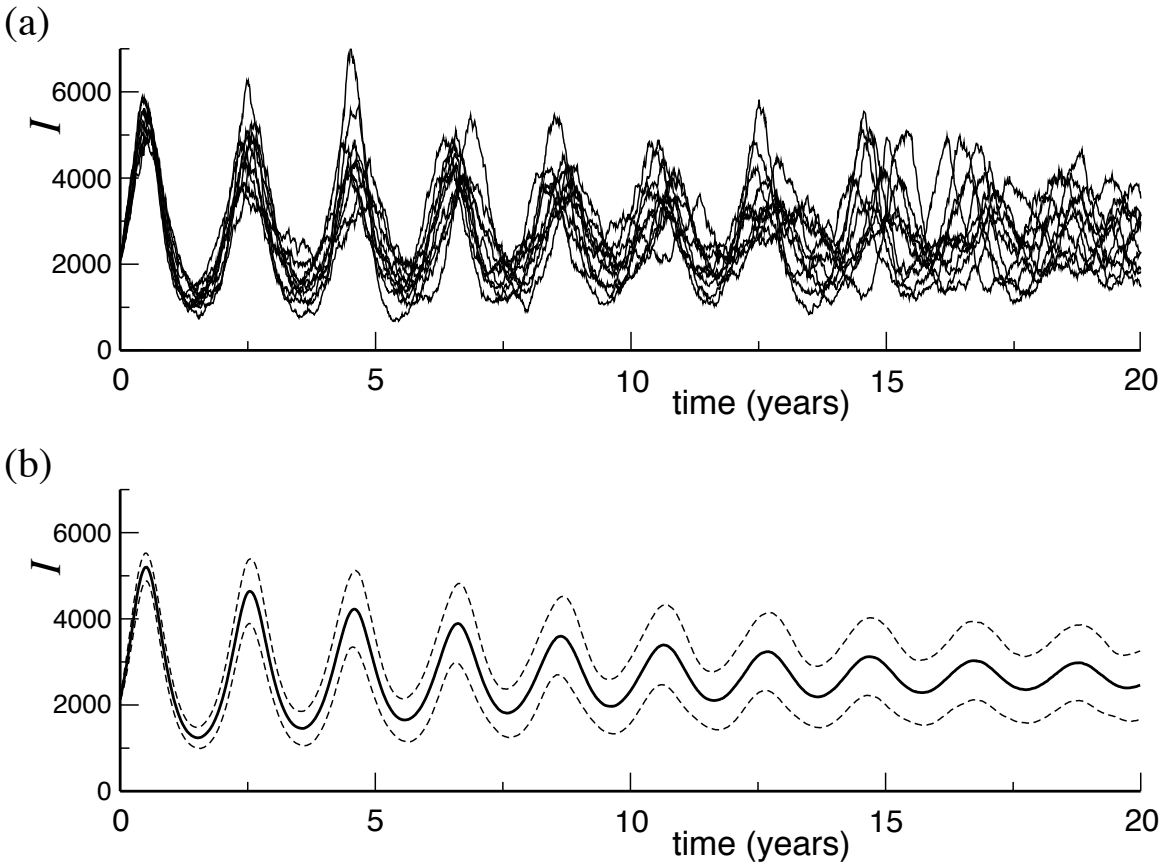
# Stochastic SIR Model: Variability Between Realizations

Individual realizations fluctuate about endemic equilibrium, drifting out of phase over time

Magnitude of fluctuations determines fadeout probability

Average exhibits damped oscillations: averaging over a collection of realizations that are drifting out of phase

Standard deviation (calculated across the set of realizations) approaches a constant

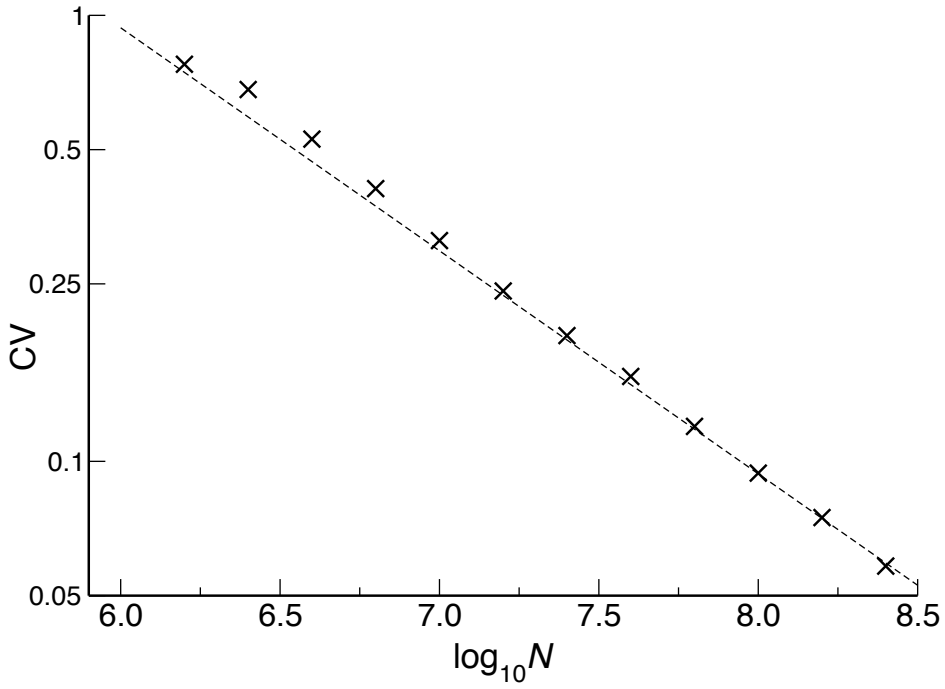



Coefficient of variation ( cv = sd/mean) is a relevant measure of variability

Lloyd (1996, PhD thesis; 2004, Theor. Popul. Biol. **65**:49)

# Stochastic SIR Model: Variability Between Realizations

Magnitude of fluctuations ( cv = sd/mean ) scales as $1/\sqrt{N}$
(central limit theorem type result)




symbols x : variability estimated from numerical simulation

Log-log plot

Dashed line has slope -1/2 (line obtained from an approximation due to Bartlett, 1956)

# Analysis: Kolmogorov Equation

$p_t(s,i)$ probability that $S(t)=s$, $I(t)=i$

Thinking about the time interval $[t, t+dt]$, how can we end up at $(s,i)$?

$$\begin{aligned} p_{t+dt}(s,i) &= \mu N dt\, p_t(s-1,i) + \mu(s+1)dt\, p_t(s+1,i) + (\beta/N)(s+1)(i-1)dt\, p_t(s+1,i-1) \\ &+ \gamma(i+1)dt\, p_t(s,i+1) + \mu(i+1)dt\, p_t(s,i+1) \\ &+ (1-\mu N dt - \mu s dt - (\beta/N)sidt - \gamma i dt - \mu i dt)p_t(s,i) \end{aligned}$$

Rearrange and let $dt \to 0$

$$\begin{aligned} \frac{d}{dt}p_t(s,i) &= \mu N\, p_t(s-1,i) + \mu(s+1)\, p_t(s+1,i) + (\beta/N)(s+1)(i-1)\, p_t(s+1,i-1) \\ &+ (\gamma+\mu)(i+1)\, p_t(s,i+1) - (\mu N - \mu s + (\beta/N)si + (\gamma+\mu)i)p_t(s,i) \end{aligned}$$

Linear equations with constant coefficients
Numerical solution feasible for small $N$, but rapidly becomes impractical

# Moment Equations

Using Kolmogorov equations, can derive equations for rates of change of expected values of functions

$$\frac{d}{dt}\mathrm{E}\left(f(S,I)\right) = \mathrm{E}\Big(\sum_{\text{events}} \text{rate of event} \times \text{change in } f \text{ due to event}\Big)$$

For example:

$$\begin{aligned}\frac{d}{dt}\mathrm{E}(I) &= \mathrm{E}\Big(\frac{\beta SI}{N} \times 1 + (\gamma+\mu)I \times (-1)\Big) \\ &= \frac{\beta}{N}\mathrm{E}(SI) - (\gamma+\mu)\mathrm{E}(I).\end{aligned}$$

Equations for 1st order moments:

$$\begin{aligned}\frac{dE(S)}{dt} &= \mu N - \mu E(S) - (\beta/N)E(SI) \\ \frac{dE(I)}{dt} &= (\beta/N)E(SI) - (\gamma+\mu)E(I)\end{aligned}$$

Observations:

1. Nonlinear transmission term leads to appearance of a 2nd order moment term
2. If $E(SI)$ were replaced by $E(S)E(I)$, we would recover the deterministic model but cannot do this because $E(SI) = E(S)E(I) + \mathrm{cov}(S,I)$

# Moment Equations

Similarly, can derive equations for 2nd order moments

Noting that $I \to I+1$ means that $I^2 \to I^2 + 2I + 1$ and $I \to I-1$ gives $I^2 \to I^2 - 2I + 1$

$$\begin{aligned}\frac{d}{dt}\mathrm{E}\left(I^2\right) &= \mathrm{E}\left(\frac{\beta SI}{N} \times (2I+1) + (\gamma+\mu)I \times (-2I+1)\right) \\ &= \frac{\beta}{N}\left(2\mathrm{E}(SI^2) + \mathrm{E}(SI)\right) + (\gamma+\mu)\left(\mathrm{E}(I) - 2\mathrm{E}(I^2)\right)\end{aligned}$$

Second order moment equations involve third order moments

Third order moment equations would involve fourth order moments, etc

For this moment equation approach to be useful, we need to truncate the set of equations at some order, requiring a moment closure approximation.

# Moment Closure

Whittle (1957): if the process was linear, realizations would follow a multivariate normal distribution… so he suggested the use of a multivariate normal approximation to close moment equations

Assume distribution of realizations at any point in time is multivariate normal (MVN)

Third order central moments of MVN distribution are zero

e.g. $E( [S-E(S)] [ I -E(I) ]^2 ) = 0$

Multiply out, gives an algebraic relationship between third order moment $E(SI^2)$ and lower order moments

Substitute into moment equation set

Limit theorems of Kurtz (1970, 1971) provide justification for this approach

# Moment Closure

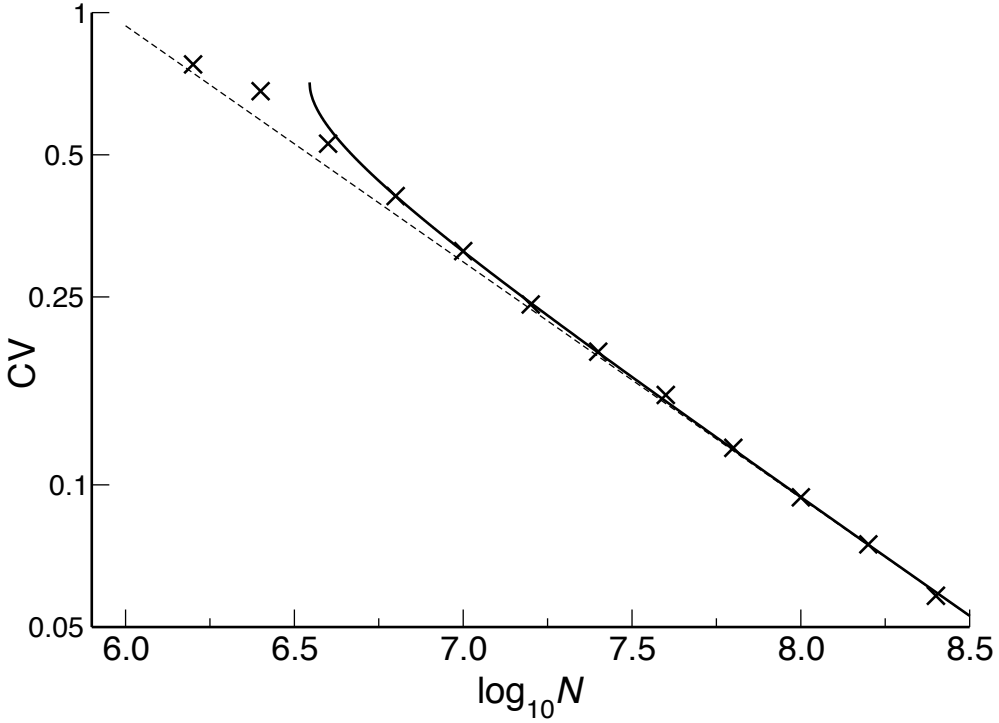
Moment equations for SIR model, in terms of variances and covariances

$$
\begin{aligned}
\frac{dE(S)}{dt} &= \mu N - \mu E(S) - (\beta/N)E(SI) \\
\frac{dE(I)}{dt} &= (\beta/N)E(SI) - (\gamma+\mu)E(I) \\
\frac{d\mathrm{Var}(S)}{dt} &= \mu N + \mu\{E(S) - 2\mathrm{Var}(S)\} + (\beta/N)\{E(SI) - 2E(I)\mathrm{Var}(S) \\
&\quad -2E(S)\mathrm{Cov}(S,I)\} \\
\frac{d\mathrm{Var}(I)}{dt} &= (\gamma+\mu)\{E(I) - 2\mathrm{Var}(I)\} + (\beta/N)\{E(SI) + 2E(S)\mathrm{Var}(I) \\
&\quad +2E(I)\mathrm{Cov}(S,I)\} \\
\frac{d\mathrm{Cov}(S,I)}{dt} &= -(\gamma+2\mu)\mathrm{Cov}(S,I) - (\beta/N)\{E(SI) - E(I)[\mathrm{Var}(S) - \mathrm{Cov}(S,I)] \\
&\quad +E(S)\,[\mathrm{Var}(I) - \mathrm{Cov}(S,I)]\}.
\end{aligned}
$$

Comment: Standard deterministic ODEs emerge by making a lower order moment closure assumption, namely 2nd order central moments are zero (set variances and covariances to zero… in other words, assume no variation between realizations)

# Moment Closure

Works well… provided population size is large enough




solid line : variability estimated from moment equations

(not shown: unstable branch of solutions)

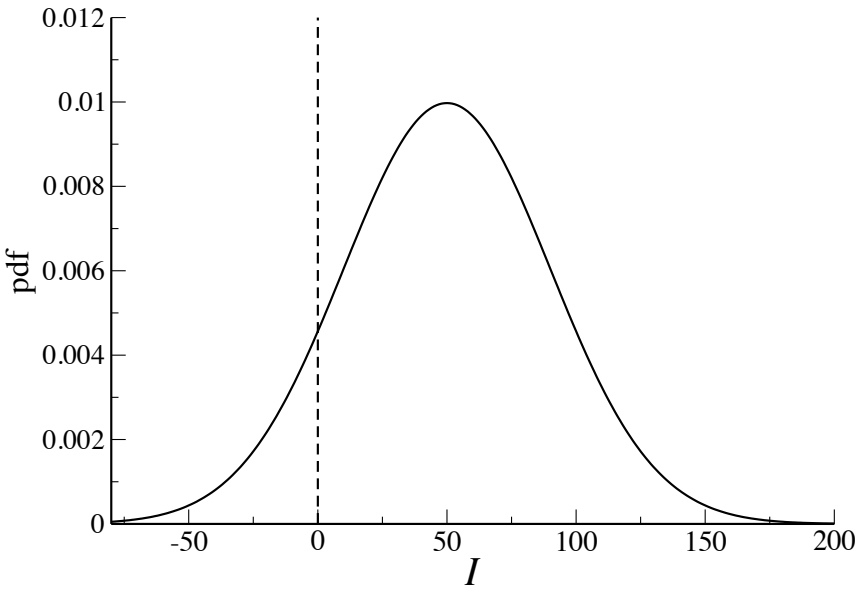
See some deviation from $N^{-0.5}$ behavior, apparently correctly captured by moment equations

Solution of moment equations blows up if *N* is too small
(Critical value of *N :* saddle node bifurcation, with stable and unstable branches colliding)

Also get divergence for larger *N* if we start too far from equilibrium

# Why does MVN break down?

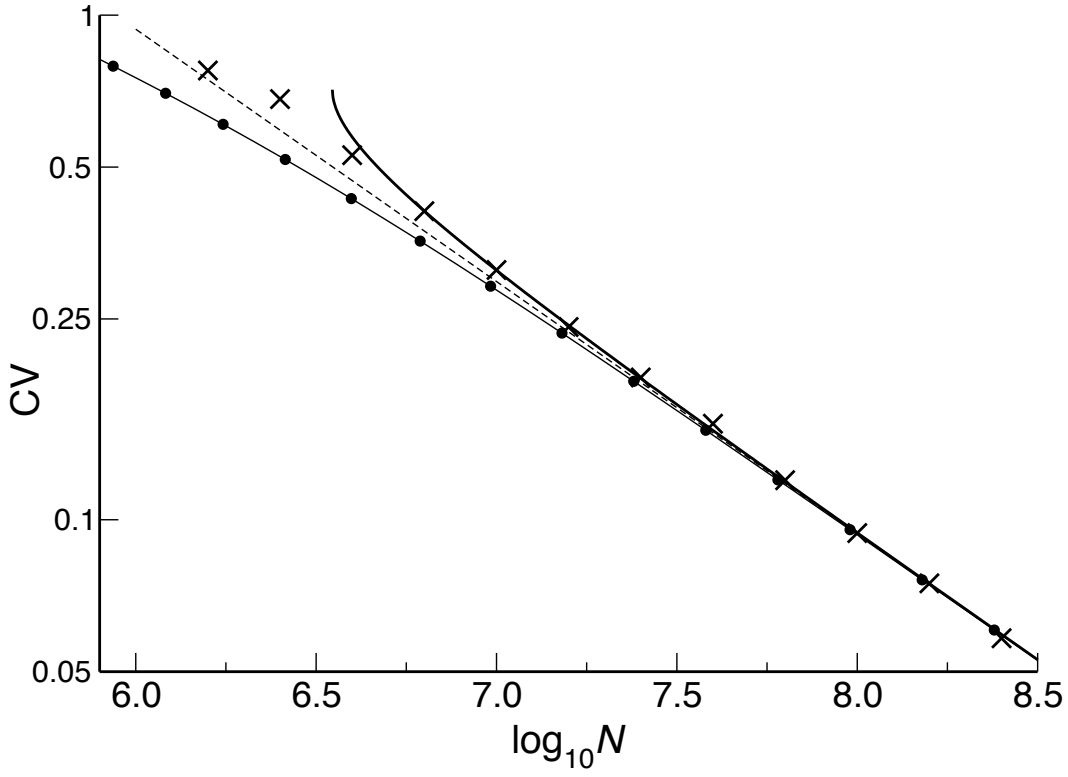
If variability is too large (cv ≈ 1), Normal distribution will have appreciable weight for negative values of *I* … but *I* must be non-negative.




Could turn to another closure approximation, e.g. Keeling's "multiplicative moment closure", which imposes relationship based on log-Normal distribution (e.g. Keeling, 2000)

# Log-Normal Moment Closure

Log-Normal moment closure approximation doesn't blow up in this case, although it does underestimate variability




Log-Normal moment closure approximation:

$$E(S^3) = \hat{V}_S^3 E(S)^3,$$

$$E(S^2 I) = \hat{V}_S \xi^2 E(S)^2 E(I),$$

$$E(S I^2) = \hat{V}_I \xi^2 E(S) E(I)^2, \quad \text{and}$$

$$E(I^3) = \hat{V}_I^3 E(I)^3.$$

where

$$E(S^2) = \hat{V}_S E(S)^2,$$

$$E(I^2) = \hat{V}_I E(I)^2,$$

$$E(SI) = \xi E(S) E(I).$$

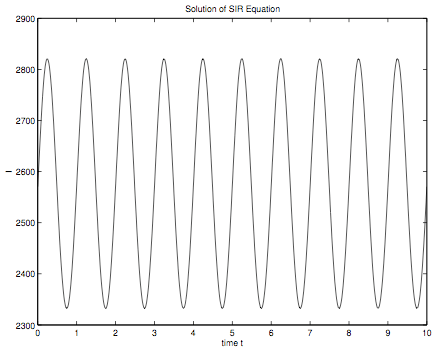
# Recap: Seasonally Forced Deterministic Model

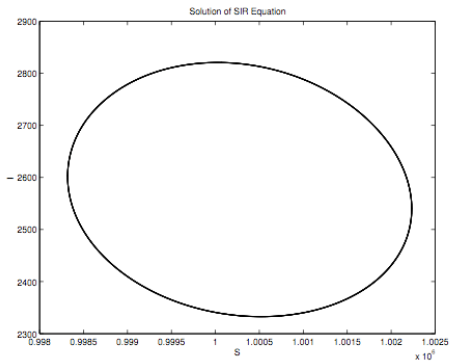
Transmission of measles has a strong seasonal component
(different levels of mixing of children during school terms and vacation)

Modeled by allowing $\beta$ to vary seasonally $\beta(t) = \beta_0\left(1 + \beta_1 \cos 2\pi t\right)$

Seasonally forced SIR model exhibits multi-annual oscillations:

Weak seasonality typically gives annual oscillations, amplitude increases with $\beta_1$




Period-doubling bifurcation seen as $\beta_1$ is further increased, giving biennial oscillations

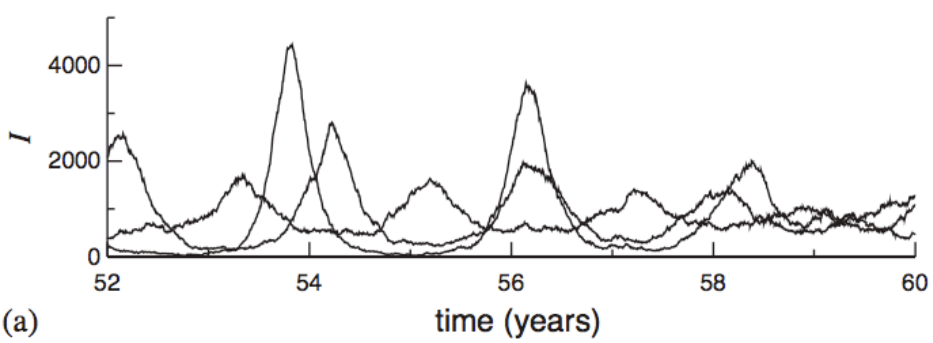
Important point: seasonality imposes a phase on solutions, unlike the non-forced model

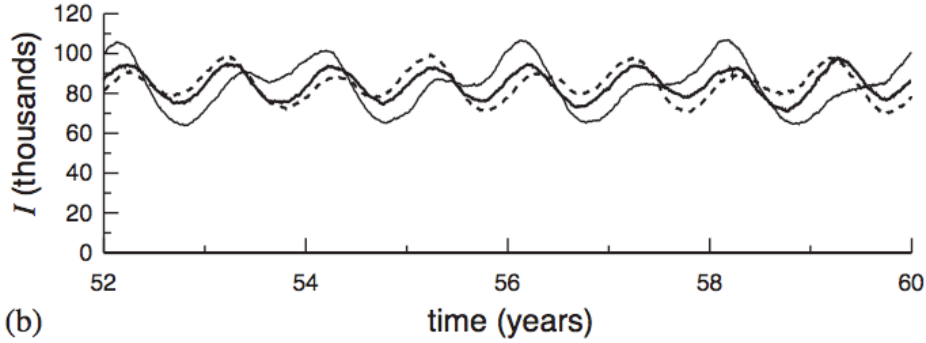
# Seasonally Forced Stochastic Model

Dynamics depends strongly on population size

e.g. three realizations from models with weak seasonality:

$N = 10^{6.5}$ $N = 10^{8.5}$




Note: the two graphs would be identical (up to scaling on vertical axis) for the deterministic model

Dynamics are a superposition of:

- deterministic annual oscillations (specific phase), amplitude independent of $N$

and

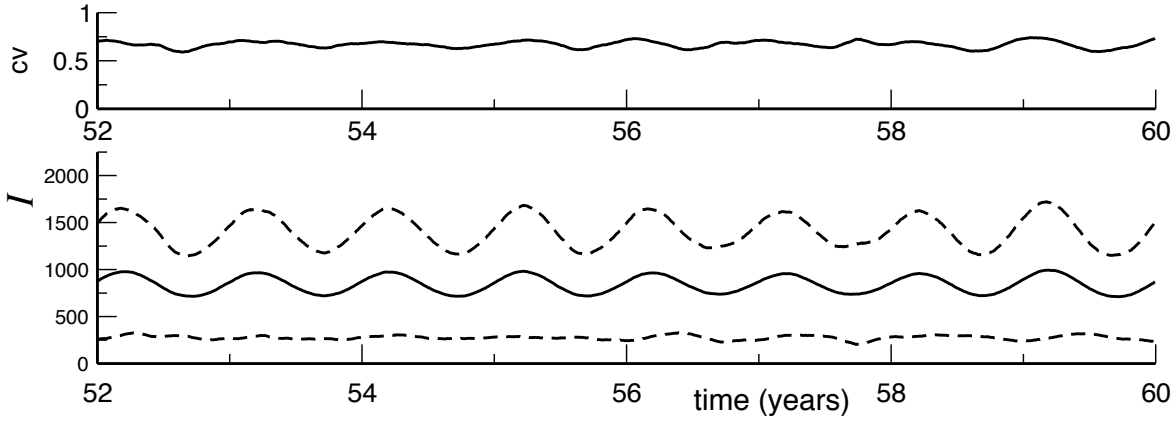
- stochastic fluctuations (no specific phase), amplitude scales as $N^{-1/2}$

# Seasonally Forced Stochastic Model: Average Behavior and Variability

- Variability changes over the course of a year
- Relative importance of deterministic and stochastic components of dynamics depends on *N*
- Deterministic dynamics more apparent in average behavior (stochastic behavior is "averaged out" when we average over a collection of out-of-phase stochastic realizations)

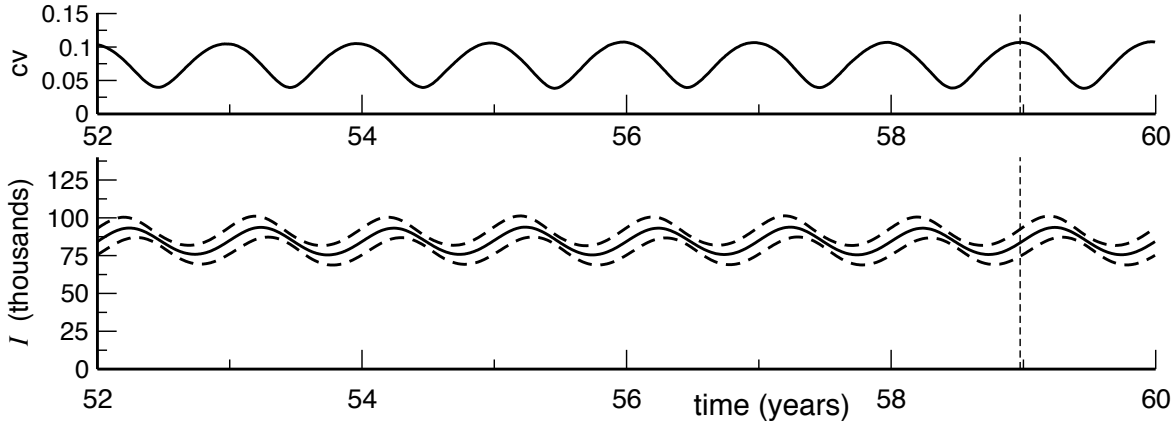
$N = 10^{6.5}$

Note: stochastic variation (envelope height) greater than deterministic variation (amplitude of oscillations in average)




$N = 10^{8.5}$

Stochastic variation (envelope height) greater than deterministic variation (amplitude of oscillations in average)



# Automating the Generation of Moment Equations

Deriving moment equations by hand is cumbersome. Fortunately, we can automate the process.

Brief outline: process involves moment generating function (MGF) $M(\theta_1, \theta_2, t) \equiv E\left(e^{\theta_1 S(t) + \theta_2 I(t)}\right)$

Important property of MGF: coefficient of $\theta_1^m \theta_2^n$ in the power series expansion of MGF is proportional to $E(S^m I^n)$.

$$M(\theta_1, \theta_2) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} {\theta_1}^j {\theta_2}^{k-j} E(S^j I^{k-j})$$

Multiply the Kolmogorov equation by exp( $\theta_1 s(t) + \theta_2 i(t)$ ) and sum over $s(t)$, $i(t)$

After some manipulation, get a PDE satisfied by MGF:

$$\frac{\partial M}{\partial t} = (\beta/N)(e^{\theta_2 - \theta_1} - 1)\frac{\partial^2 M}{\partial \theta_1 \partial \theta_2} + \mu(e^{-\theta_1} - 1)\frac{\partial M}{\partial \theta_1} + (\gamma + \mu)(e^{-\theta_2} - 1)\frac{\partial M}{\partial \theta_2} + \mu N(e^{\theta_1} - 1)M$$

Expand both sides of PDE as a power series in $\theta_1$ and $\theta_2$, then equating coefficients of $\theta_1^m \theta_2^n$ on both sides gives a set of ODEs for the time evolution of the moments

# Automating the Generation of Moment Equations, II

Application of MVN approximation is easier if we work with the cumulant generating function, $K(\theta_1, \theta_2, t) \equiv \ln M(\theta_1, \theta_2, t)$, because the cumulant generating function of a MVN is a quadratic:

$$K(\theta_1, \theta_2) = \theta_1 E(S) + \theta_2 E(I) + \tfrac{1}{2}\theta_1^2 \mathrm{Var}(S) + \theta_1\theta_2 \mathrm{cov}(S,I) + \tfrac{1}{2}\theta_2^2 \mathrm{Var}(I)$$

Using Maple, it's easy to substitute $M = \exp(K)$ into the PDE for $M$, expand and equate coefficients

Overkill for doing the SIR model, but for a general population model, it's easy to convert transition information (rate of transition and resulting changes in state variables) for a general term that occurs at rate $\alpha X_1^{k_1} X_2^{k_2} \cdots X_m^{k_m}$ and involves changes $X_i \to X_i + \delta_i$ into an appropriate term in the PDE:

$$\alpha \left( e^{\Sigma \delta_i \theta_i} - 1 \right) \frac{\partial^{k_1 + \cdots + k_m} M}{\partial \theta_1^{k_1} \cdots \partial \theta_m^{k_m}}$$

So it's trivial to construct the PDE...

… and once Maple has expanded and equated coefficients, it will even generate the Matlab code (and LaTeX code)

see also C.S. Gillespie (2009) IET Syst. Biol. **3**:52.

# Moment Equations for Higher Dimension Systems

*m* state variables means there are $m(m+3)/2$ 1st and 2nd order moment equations

Automation facilitates use of moment equations for more complex models!

e.g. SEIR model (3 state variables), $SI^nR$ ($n$+1 state variables), $n$-patch SIR model ($2n$), vector-host model (e.g. mosquito-borne infections)

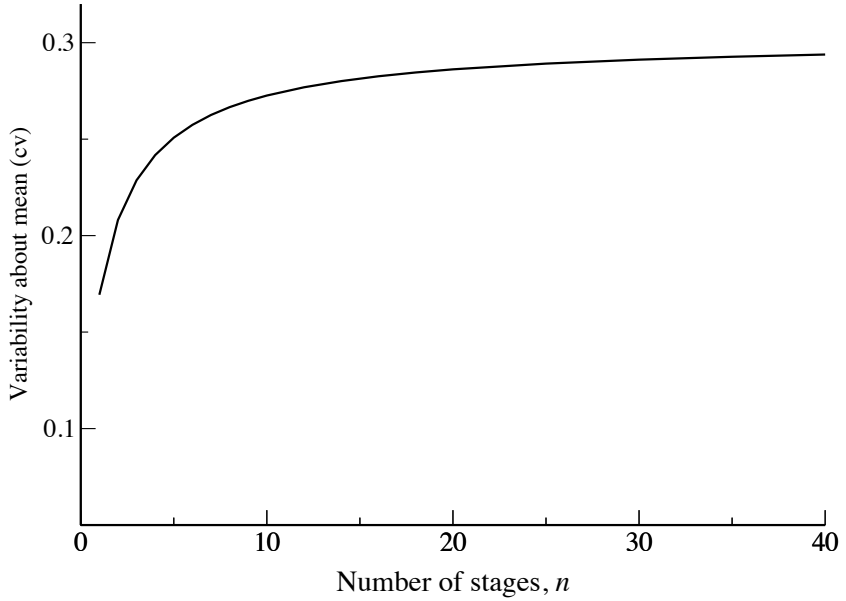
SIR model with $n$-stage infectious period

(models less variable infectious periods compared to exponential--- more realistic)


Less variable infectious period increases variability, hence decreases persistence

(Lloyd, Theor. Popul. Biol., 2001)

(902 equations in moment set for 40-stage model)



# Origin of the Sustained Fluctuations in the SIR model

Sustained fluctuations in the stochastic SIR model have long been known, as has the basic idea of why they occur:

Resonance effect between intrinsic damped oscillatory behavior of the deterministic model and forcing due to demographic stochasticity

Two recent papers have made this idea more concrete:

Kuske, Gordillo and Greenwood (2007) "Sustained oscillations via coherence resonance in SIR"

Alonso, McKane and Pascual (2007) "Stochastic amplification in epidemics"

Employed various approximations to calculate the power spectrum of fluctuations,

demonstrating the resonant response of the deterministic system to (essentially) white noise forcing of demographic stochasticity

McKane's group has carried out similar analyses for a number of SIR-type models, including the *n*-stage model (Black et al, 2009)

# Why is Measles so Susceptible to Demographic Stochasticity?

Key epidemiological points

- recovery from infection confers permanent immunity
- infectious period is short
- highly infectious (before vaccination, almost everyone caught measles)

Separation of timescales between short duration of infection (~ one week) and slow turnover of susceptible population (lifespan of ~75 years)

The short duration of infection means that only a small fraction of the population is infected at any given time ( roughly 1/(52*75) ), so even if *N* is large, the number of infectives is typically small.

# Kurtz's Results

Results from Kurtz (1970) relates behavior of stochastic model to that of ODE

Considers a family of Markov chains, $\mathbf{X}_n(t)$, indexed by parameter *n* (think: total population size) that are "density dependent population processes", i.e. for which transition rates between states **x** and **x+k** have the form

$$q_{\mathbf{x},\mathbf{x}+\mathbf{k}} = nf(\mathbf{x}/n, \mathbf{k})$$

Defines $\mathbf{F}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{k} f(\mathbf{x}, \mathbf{k})$

Require $f_{\mathbf{m}}$ and **F** to satisfy some regularity conditions in an open set *E*

$$|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})| \leq M_E |\mathbf{x} - \mathbf{y}|$$

$$\sup_{\mathbf{x}\in E} \sum_{\mathbf{k}} |\mathbf{k}| f(\mathbf{x}, \mathbf{k}) < \infty \qquad \lim_{d\to\infty} \sup_{\mathbf{x}\in E} \sum_{|\mathbf{k}|>\mathbf{d}} |\mathbf{k}| f(\mathbf{x}, \mathbf{k}) = 0$$

# Kurtz's Results

Then: if we have a deterministic trajectory $\mathbf{X}(s,\mathbf{x}_0)$ that starts at $\mathbf{x}_0$ , governed by the ODE $d\mathbf{X}/dt = \mathbf{F}(\mathbf{X})$, that stays in the open set *E* at least until time *t* ,

then $\lim_{n\to\infty} (1/n)\mathbf{X}_n(0) = \mathbf{x}_0$ implies, for every $\delta > 0$

$$\lim_{n\to\infty} \mathrm{Prob}\left\{ \sup_{s\le t} \left| \frac{1}{n}\mathbf{X}_n(s) - \mathbf{X}(s,\mathbf{x}_0) \right| > \delta \right\} = 0$$

In the large *n* limit, trajectories of the stochastic model stay close to the deterministic trajectory

Doesn't say how large *n* has to be. Kurtz 1971 goes further and shows that fluctuations about the deterministic model converges (in some sense) to a diffusion process

# Diffusion Approximations to Variation About Endemic State

An alternative approach to the moment equations uses a diffusion approximation to estimate variation about the endemic state

Idea is to linearize the system about the endemic equilibrium. Fluctuations are then governed by the sum of a linear drift term (towards equilibrium) and a diffusion term

Gives a multivariate Ornstein-Uhlenbeck process for the fluctuations; a process whose limiting probability distribution is known, and whose variance-covariance matrix satisfies

$$A\Sigma+\Sigma A^T = -V$$

Here, A is the linear matrix describing the drift (Jacobian matrix of system, evaluated at equilibrium), V is the constant diffusion matrix

Variability reflects balance between random effects and restoring force of equilibrium

# Practical Exercise 1

Simulate the stochastic SIR model (no demography) and generate time series showing trajectories of $I(t)$.

Take $\gamma = 1$, different values of $R_0 = \beta$ [think what might be appropriate/interesting values to examine], and run each realization until the epidemic ends. I suggest a population size of $N = 1000$ and one initial infective.

Collect 1000 realizations. Plot 5 or 10 realizations and the outbreak size distribution. Estimate the probability of a major outbreak from your outbreak size distribution. Compare your answer to that predicted by the branching process theory.

Repeat for a larger initial number of infectives (3, 5 or 10).

Something interesting to try would be to compare the corresponding deterministic model to the average behavior of the stochastic model and to the average of the stochastic model *conditioned on non-extinction* (i.e. take averages over those realizations for which $I(t)>0$ at each point in time)

# Practical Exercise 2

Write code to simulate a branching process model for Poisson offspring distribution with mean $R_0$. Start from one initial individual and allow a realization to continue until the number of individuals in a generation reaches 100.

(Hint to speed up your code: what is the sum of *n* independent Poisson distributions?)

Show 5 or 10 realizations of the process and the mean behavior across 1000 realizations. Numerically estimate the probability of extinction of the process. Calculate this extinction probability using branching process theory.

Modify your code to simulate the branching process with geometric offspring distribution. (Hint: what is the sum of *n* independent geometric distributions?)

Compare the behavior of the two processes, specifically the probability of extinction.