

Introduction to SDE Models in Population Biology

Lecture 4: Long-Term Dynamics and Research Directions

2-hour PhD short course – 50-page beginner version

Learning objectives and two-hour plan

Students should understand

- the Markov semigroup and infinitesimal generator at a working level;
- invariant measures as stochastic analogues of long-term attractors;
- ergodic averages and their biological meaning;
- how boundary invariant measures lead to stochastic persistence criteria;
- how the elementary SDE toolkit connects with current population-biology research.

Suggested timing: 0–25 min Markov semigroups and generators; 25–50 min invariant measures; 50–60 min break; 60–85 min ergodicity; 85–110 min persistence through boundary growth rates; 110–120 min research directions and projects.

Textbook guide and transition to research literature

From Øksendal we use the Markov and generator viewpoint for diffusions. From Mao we use Lyapunov methods, recurrence/stability ideas, and long-time estimates. The conceptual transition is:

sample-path SDE calculations



probability laws and invariant measures



persistence in multispecies systems

The final part of the lecture explains the ideas students should recognize when reading modern papers, without attempting a full general persistence theorem.

The Markov semigroup

Let X_t^x denote the solution started from $X_0 = x$. Define

$$P_t\varphi(x) = \mathbb{E}_x[\varphi(X_t)].$$

Then $(P_t)_{t \geq 0}$ is a Markov semigroup under standard assumptions:

$$P_0\varphi = \varphi, \quad P_{t+s}\varphi = P_t(P_s\varphi).$$

- P_t evolves observables forward in time.
- The dual semigroup evolves probability distributions.
- The Markov property says the future depends on the present state, not on the entire past history.

Transition probabilities and the Markov property

For a Markov process X_t , define the transition kernel

$$P_t(x, A) = \mathbb{P}_x\{X_t \in A\}.$$

The Markov property says that, conditional on the present state, the future is independent of the past. The semigroup acts on bounded measurable functions by

$$P_t f(x) = \int f(y) P_t(x, dy) = \mathbb{E}_x[f(X_t)].$$

The Chapman–Kolmogorov property becomes

$$P_{t+s} = P_t P_s.$$

Feller property and why analysts care

A Markov semigroup is Feller if

$$f \in C_b(E) \implies P_t f \in C_b(E).$$

For SDEs with continuous coefficients and suitable well-posedness, the Feller property often follows from continuous dependence on initial data. Feller continuity is useful for:

- proving existence of invariant measures from tight occupation measures;
- passing to limits in weak convergence arguments;
- connecting probabilistic dynamics with semigroup methods familiar from PDEs.

Generator and Dynkin's formula

For

$$dX = b(X)dt + \Sigma(X)dW, \quad a = \Sigma\Sigma^\top,$$

the generator is

$$L\varphi = b \cdot \nabla\varphi + \frac{1}{2} \text{tr}(aD^2\varphi).$$

For suitable stopping times τ ,

$$\mathbb{E}_x[\varphi(X_{t\wedge\tau})] = \varphi(x) + \mathbb{E}_x \int_0^{t\wedge\tau} L\varphi(X_s)ds.$$

This is Dynkin's formula. **Why useful?** It converts pathwise stochastic dynamics into expected drift identities and is the basis of many Lyapunov recurrence arguments.

Generator as the derivative of the semigroup

Formally, the generator satisfies

$$Lf = \lim_{t \downarrow 0} \frac{P_t f - f}{t}.$$

For a diffusion

$$dX = b(X)dt + G(X)dW,$$

this becomes

$$Lf = \nabla f^T b + \frac{1}{2} \text{tr}(GG^T D^2 f).$$

The generator is therefore the infinitesimal deterministic operator associated with the stochastic dynamics. It is the bridge between SDEs, PDEs, martingales, and Lyapunov methods.

Dynkin formula from the martingale decomposition

Itô's formula gives

$$f(X_t) = f(X_0) + \int_0^t Lf(X_s)ds + M_t,$$

where M_t is a local martingale. Taking expectations yields Dynkin's formula:

$$\mathbb{E}_x f(X_t) - f(x) = \mathbb{E}_x \int_0^t Lf(X_s)ds.$$

Thus the sign of LV controls the average evolution of a Lyapunov function V . This is the stochastic analogue of differentiating V along trajectories of an ODE.

Backward and forward Kolmogorov viewpoints

Formally, if

$$u(t, x) = P_t \varphi(x),$$

then u satisfies the backward equation

$$\partial_t u = Lu, \quad u(0, x) = \varphi(x).$$

If the law of X_t has density $p(t, x)$, then p satisfies the forward/Fokker–Planck equation

$$\partial_t p = L^* p.$$

Bridge to the user's background

This is where stochastic population dynamics meets parabolic PDE theory: the SDE evolves sample paths, while the Fokker–Planck equation evolves probability densities.

Backward Kolmogorov equation

Let

$$u(t, x) = \mathbb{E}_x[\varphi(X_t)] = P_t\varphi(x).$$

Under sufficient regularity,

$$\partial_t u = Lu, \quad u(0, x) = \varphi(x).$$

This is the backward Kolmogorov equation. The state variable x is the initial condition, so the equation describes how expected future observables change as a function of the starting state.

Forward equation and probability densities

If X_t has a density $p(t, x)$, then formally

$$\partial_t p = L^* p,$$

where L^* is the adjoint of the generator. In one dimension,

$$\partial_t p = -\partial_x(bp) + \frac{1}{2}\partial_{xx}(\sigma^2 p).$$

A stationary density solves

$$L^* p = 0.$$

This PDE viewpoint complements the probabilistic invariant-measure viewpoint.

Invariant probability measures

A probability measure μ is invariant if

$$\int P_t \varphi(x) \mu(dx) = \int \varphi(x) \mu(dx)$$

for all suitable φ and $t \geq 0$. Equivalently, if $X_0 \sim \mu$, then

$$X_t \sim \mu \quad \text{for every } t \geq 0.$$

Biological interpretation: the population still fluctuates, but its statistical distribution is unchanged in time.

Invariance written in three equivalent ways

A probability measure π is invariant if

$$\pi P_t = \pi.$$

Equivalently, for every bounded measurable f ,

$$\int P_t f \, d\pi = \int f \, d\pi.$$

For sufficiently regular f in the generator domain, invariance implies

$$\int Lf \, d\pi = 0.$$

The last identity is often convenient for deriving moment relations under an invariant law.

Occupation measures along one trajectory

Define the empirical occupation measure

$$\Pi_t = \frac{1}{t} \int_0^t \delta_{X_s} ds.$$

For a test function f ,

$$\int f d\Pi_t = \frac{1}{t} \int_0^t f(X_s) ds.$$

Thus long-time averages are exactly integrals against the occupation measure. If a subsequence Π_{t_n} converges weakly and the process is Feller, the limit is often an invariant probability measure.

How invariant measures are often obtained

A common strategy is the Krylov–Bogoliubov idea. Consider the empirical occupation measures

$$\Pi_T = \frac{1}{T} \int_0^T \delta_{X_t} dt.$$

If the family $\{\Pi_T\}$ is tight, then weakly convergent subsequences may exist. Under suitable Feller/continuity assumptions, a weak limit is invariant. So a standard proof pattern is

Lyapunov drift condition $\Rightarrow$ tight occupation measures
 $\Rightarrow$ existence of an invariant measure

Krylov–Bogoliubov averaging

Starting from an initial distribution ν , define

$$\mu_T = \frac{1}{T} \int_0^T \nu P_t dt.$$

If the family $\{\mu_T : T \geq 1\}$ is tight, then Prokhorov's theorem gives weakly convergent subsequences. For a Feller process, any weak limit of such a subsequence is invariant. Hence the central technical

issue is often not solving $L^*p = 0$ explicitly, but proving enough recurrence or moment control to obtain tightness.

Tightness from a coercive Lyapunov function

Suppose $V(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ and

$$LV(x) \leq C - cV(x)$$

for positive constants C, c . Then Dynkin's formula gives a uniform bound on time averages of $V(X_t)$. Since sublevel sets $\{V \leq R\}$ are compact, Markov's inequality implies that occupation measures put arbitrarily large mass in a compact set when R is large. This is the basic route from a drift condition to tightness.

A typical Lyapunov drift condition for recurrence

Suppose there is a coercive $V \geq 1$ and constants $c, d > 0$ such that outside a compact set,

$$LV(x) \leq -cV(x) + d.$$

Then large values of V have negative expected drift.

- This discourages escape to infinity.
- It yields moment bounds and tightness of long-time occupation measures.
- With additional irreducibility conditions, it can help prove uniqueness and ergodicity.

This stochastic Foster–Lyapunov philosophy is one of the most reusable ideas in the subject.

Foster–Lyapunov drift condition

A frequently used recurrence condition is

$$LV \leq -c \quad \text{outside a compact set } K.$$

A stronger geometric form is

$$LV \leq -cV + d\mathbf{1}_K.$$

Such inequalities say that the stochastic dynamics has an average restoring drift toward a bounded region. Combined with an irreducibility or minorization condition, they can yield positive recurrence, uniqueness of the invariant law, and quantitative ergodicity.

Irreducibility and accessibility

A drift condition alone may not guarantee a unique invariant distribution. The state space can split into dynamically disconnected pieces. Irreducibility-type assumptions ensure that the process can move between sufficiently many regions of state space. For nondegenerate diffusions this may follow from ellipticity. For degenerate population systems it may require support theorems, controllability ideas, or model-specific arguments.

Logistic diffusion as the running example

Recall

$$dX = X(a - bX)dt + \sigma X dW, \quad a > \frac{\sigma^2}{2}.$$

Its invariant density is

$$p(x) = \frac{r^k}{\Gamma(k)} x^{k-1} e^{-rx}, \quad k = \frac{2a}{\sigma^2} - 1, \quad r = \frac{2b}{\sigma^2}.$$

The stationary mean is

$$\int_0^\infty xp(x)dx = \frac{a - \sigma^2/2}{b}.$$

This scalar model illustrates the full chain: corrected growth rate $\rightarrow$ persistence $\rightarrow$ invariant distribution $\rightarrow$ ergodic average.

Ergodicity: long trajectories reveal stationary averages

If μ is an ergodic invariant measure and suitable hypotheses hold, then for integrable φ ,

$$\frac{1}{T} \int_0^T \varphi(X_t) dt \longrightarrow \int \varphi d\mu \quad \text{a.s.}$$

Examples of biologically useful observables:

- $\varphi(x) = x_i$: average abundance of species i ;
- $\varphi(x) = \mathbf{1}_{\{x_i > c\}}$: fraction of time abundance exceeds a threshold;
- $\varphi(x) = x_i x_j$: long-term interaction statistic;
- epidemic models: fraction of time prevalence exceeds a public-health threshold.

Ergodic theorem for stationary Markov processes

If π is an ergodic invariant probability measure and f is integrable, then for a stationary realization,

$$\frac{1}{t} \int_0^t f(X_s) ds \rightarrow \int f d\pi \quad \text{a.s.}$$

Under positive recurrence and suitable irreducibility, the same limit can hold from arbitrary initial states. This is why a single sufficiently long time series can estimate stationary biological averages.

Invariant is not the same as ergodic

An invariant measure may be a mixture of different long-term regimes. An ergodic invariant measure is an “indecomposable” stationary regime: time averages along almost every stationary trajectory equal expectations under that same measure.

Deterministic analogy

A deterministic system can have several invariant sets. Similarly, a stochastic system can have several invariant measures, especially when different boundary communities are dynamically possible.

This distinction becomes essential in multispecies persistence theory.

The extinction boundary

For n populations with state $x \in \mathbb{R}_+^n$, define

$$\partial\mathbb{R}_+^n = \{x : \text{at least one } x_i = 0\}.$$

Different parts of the boundary represent different reduced communities:

- all populations extinct;
- one surviving resident species;
- a subset of interacting species;
- disease-free dynamics in epidemic models.

A modern persistence theory studies invariant probability measures supported on these boundary subsystems.

Average growth rate of a missing population

Suppose population i satisfies

$$dX_i = X_i \left[f_i(X) dt + \sum_j \sigma_{ij}(X) dW_j \right].$$

Its corrected per-capita drift is

$$g_i(x) = f_i(x) - \frac{1}{2} \sum_j \sigma_{ij}^2(x).$$

For an invariant measure μ supported on a boundary where species i is absent, define

$$\lambda_i(\mu) = \int g_i(x) \mu(dx).$$

This is the average invasion rate of species i against the resident stochastic environment described by μ .

Boundary invariant measures in population systems

Because coordinate faces are often invariant, the boundary may support several invariant measures:

- the extinction measure δ_0 ;
- single-species stationary laws;
- stationary laws of two-species subsystems;
- mixtures or more complicated ergodic measures on boundary faces.

A general stochastic persistence theorem must examine the growth of missing species against every *relevant ergodic boundary measure*, not only one obvious resident state.

General invasion rate against a boundary measure

For

$$dX_i = X_i f_i(X) dt + X_i g_i(X) dW_i,$$

the corrected per-capita growth is

$$q_i(x) = f_i(x) - \frac{1}{2} g_i^2(x).$$

If μ is an invariant probability measure with species i absent, define

$$\lambda_i(\mu) = \int q_i(x) \mu(dx).$$

This quantity is the average exponential growth rate of species i when introduced at very low density into the resident community described by μ .

A schematic persistence criterion

A common form of stochastic persistence condition is: *for every ergodic invariant measure μ supported on the extinction boundary, at least one missing population has positive average growth rate.* More flexible theorems use positive weights p_i and require

$$\sum_i p_i \lambda_i(\mu) > 0$$

for every relevant boundary invariant measure μ .

Interpretation

Every boundary community is invasible in a weighted average sense, so the full process is repelled from extinction states over long time scales.

A weighted invasion condition

In multispecies persistence theorems one often seeks positive weights p_i such that

$$\sum_{i=1}^n p_i \lambda_i(\mu) > 0$$

for every ergodic invariant measure μ supported on the extinction boundary. This means that every boundary community has, in a weighted sense, at least one direction of positive growth back toward the interior. The exact hypotheses and conclusion vary among theorems, but this weighted condition is a recurring structural idea.

Stochastic persistence via occupation measures

One formulation of persistence is that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\liminf_{t \rightarrow \infty} \Pi_t \{x : \min_i x_i \geq \delta\} \geq 1 - \varepsilon$$

with high probability or almost surely, depending on the theorem. This says that the process spends asymptotically little time near extinction faces. The occupation-measure viewpoint is particularly natural because invasion rates are themselves averages under invariant measures.

Competition example in invariant-measure language

For the two-species competition system, the X -only stationary law μ_X satisfies

$$\int x \mu_X(dx) = \frac{a_1 - \sigma_1^2/2}{b_{11}}.$$

The missing species Y has invasion rate

$$\lambda_Y(\mu_X) = a_2 - \frac{\sigma_2^2}{2} - b_{21} \int x \mu_X(dx).$$

Thus

$$\lambda_Y(\mu_X) = a_2 - \frac{\sigma_2^2}{2} - b_{21} \frac{a_1 - \sigma_1^2/2}{b_{11}},$$

which recovers the formula from Lecture 3 in a general framework.

What must accompany positive invasion rates?

Positive boundary growth rates are powerful, but rigorous persistence results usually also need hypotheses ensuring:

- global well-posedness and positivity;
- tightness or boundedness in probability;
- enough continuity/Feller structure;
- control of degenerate noise and accessibility of state-space regions;
- a precise persistence conclusion: occupation measures away from the boundary, persistence in probability, or existence of an interior invariant measure.

Lesson for reading papers: always separate the growth-rate condition from the technical assumptions that make it effective.

Environmental versus demographic noise

Environmental noise

- acts on per-capita rates;
- often multiplicative in abundance;
- may be correlated across species;
- often represented by Brownian forcing.

Demographic noise

- originates from random births/deaths;
- is strongest at small population sizes;
- often begins with a birth–death or jump process;
- diffusion approximations may have different noise scaling.

The biological mechanism should determine the stochastic term, not mathematical convenience alone.

Beyond autonomous diffusion models

Natural next directions include:

- **Regime switching:** coefficients depend on a Markov chain representing environmental states;
- **Seasonal forcing:** time-periodic or nonautonomous stochastic systems;
- **Spatial patch models:** coupled SDEs with migration among habitats;
- **Stochastic reaction–diffusion equations:** space and environmental noise together;
- **Jump noise:** catastrophes, harvesting shocks, sudden outbreaks;
- **Parameter inference:** connecting SDEs with biological time-series data.

Model extensions: switching, jumps, delay, and space

The diffusion framework is only one starting point. Population models may include:

- **regime switching**: coefficients depend on a finite-state Markov chain;
- **jump noise**: rare catastrophes or pulses modeled by Poisson jumps;
- **delay**: maturation or gestation creates functional SDEs;
- **space**: stochastic reaction–diffusion equations or stochastic patch systems.

The same themes persist: positivity, invariant objects, Lyapunov functions, logarithmic growth rates, and boundary dynamics.

How to read a stochastic population-biology paper

For each paper, identify:

- ① the biological variables and deterministic baseline model;
- ② the source and mathematical form of randomness;
- ③ the state space and extinction boundary;
- ④ the main long-time conclusion: extinction, persistence, invariant measure, ergodicity;
- ⑤ the principal proof tool: log transform, comparison, Lyapunov function, boundary growth rate, coupling, PDE method;
- ⑥ which assumptions are biological and which are technical.

This checklist is more useful than trying to memorize individual theorems.

Suggested final mini-project

Each student chooses a stochastic population model related to their research interests and prepares a short analysis:

- derive the SDE from a deterministic model and explain the noise term;
- prove or justify positivity/global existence;
- compute one logarithmic growth rate or invasion rate;
- identify candidate boundary invariant measures;
- state one rigorous theorem from the literature and outline its proof strategy;
- propose one extension relevant to population biology.

A 10–12 minute presentation is enough if the mathematical mechanism is explained clearly.

Beginner review: what is a Markov process?

A process X_t is Markov if, given the present state, the future does not depend on the past:

$$\mathbb{P}(X_{t+s} \in A \mid \mathcal{F}_t) = \mathbb{P}(X_{t+s} \in A \mid X_t).$$

For an SDE with time-independent coefficients,

$$dX_t = b(X_t)dt + \Sigma(X_t)dW_t,$$

the solution is typically a Markov process.

Interpretation

The current population sizes contain all information needed to determine the probability law of future population sizes.

Semigroup analogy with deterministic flows

For an ODE $\dot{x} = f(x)$, the solution map is

$$x \mapsto \Phi_t(x).$$

For an SDE, the future state is random, so we act on test functions:

$$P_t f(x) = \mathbb{E}_x[f(X_t)].$$

This family $(P_t)_{t \geq 0}$ is the Markov semigroup. The semigroup property

$$P_{t+s} = P_t P_s$$

is the stochastic analogue of

$$\Phi_{t+s} = \Phi_t \circ \Phi_s.$$

Generator calculation recipe

For a diffusion

$$dX_t = b(X_t)dt + \Sigma(X_t)dW_t,$$

let

$$a(x) = \Sigma(x)\Sigma^T(x).$$

For a smooth test function f , the generator is

$$Lf(x) = \sum_i b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

Beginner recipe

First compute all first derivatives, then all second derivatives, then insert the drift vector and covariance matrix.

Dynkin formula example for the logistic diffusion

For

$$dX_t = X_t(a - cX_t)dt + \sigma X_t dW_t,$$

choose $f(x) = x$. Since $f'(x) = 1$ and $f''(x) = 0$,

$$Lf(x) = x(a - cx).$$

Dynkin's formula gives

$$\mathbb{E}X_t = X_0 + \int_0^t \mathbb{E}[X_s(a - cX_s)]ds.$$

This is not closed because $\mathbb{E}X_s^2$ appears. The first moment depends on the second moment. This illustrates why invariant measures and Lyapunov methods are often more useful than moment equations alone.

Invariant measure intuition: repeated observation

Suppose a population model has an invariant probability measure π . If X_0 is distributed according to π , then

$$X_t \sim \pi \quad \text{for all } t \geq 0.$$

In applications, π describes the long-run distribution of the population under continuing environmental variability. Instead of converging to a deterministic equilibrium, the population may fluctuate forever around a statistical equilibrium.

Biological interpretation

The invariant measure replaces the deterministic equilibrium point as the object describing long-term behavior.

Tightness intuition through a simple inequality

A family of probability measures $\{\mu_t\}$ on $\mathbb{R}_+$ is tight if most of its mass stays in a bounded interval. If there is a function $V(x) \rightarrow \infty$ as $x \rightarrow \infty$ and

$$\sup_t \int V(x) \mu_t(dx) < \infty,$$

then Markov's inequality gives

$$\mu_t(\{x : V(x) > R\}) \leq \frac{1}{R} \int V(x) \mu_t(dx).$$

For large R , this is small uniformly in t . Thus Lyapunov bounds imply tightness of occupation measures.

Ergodicity by analogy

For a deterministic stable equilibrium x^* ,

$$\lim_{t \rightarrow \infty} X_t = x^*.$$

For an ergodic stochastic system with invariant measure π ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt = \int f(x) \pi(dx) \quad \text{a.s.}$$

Thus a single long trajectory reveals ensemble averages. This is why simulation of one sufficiently long stochastic trajectory can estimate stationary means when ergodicity is known.

Boundary invariant measures: simple examples

In a two-species model, possible boundary invariant measures include:

- $\delta_{(0,0)}$, representing extinction of both species;
- π_X , supported on the X -axis, representing persistence of X alone;
- π_Y , supported on the Y -axis, representing persistence of Y alone.

Persistence theory studies whether the full system tends to move away from all such boundary measures. For each boundary measure, one computes the average growth rate of the missing species or species group.

Persistence theorem recipe for beginners

A typical stochastic persistence proof has the following structure.

- 1 Prove positivity and non-explosion.
- 2 Identify all relevant boundary invariant measures.
- 3 Compute invasion rates of missing species against those measures.
- 4 Find positive weights so that a weighted sum of invasion rates is positive on the boundary.
- 5 Use a logarithmic Lyapunov function to show the occupation measure spends little time near the boundary.

The details can be technical, but the conceptual structure is close to deterministic average-Lyapunov persistence theory.

A beginner reading plan for research papers

When reading a stochastic population-biology paper, first identify:

- 1 the state space and boundary;
- 2 the drift terms and noise terms;
- 3 which noise is environmental and which is demographic;
- 4 the positivity and boundedness argument;
- 5 the boundary invariant measures;
- 6 the invasion rates or principal exponents;
- 7 the final persistence or extinction conclusion.

Only after this overview should one study the technical estimates line by line.

Four recurring principles

- Itô correction changes biologically meaningful growth thresholds.
- Positivity and non-explosion are part of model validity, not technical afterthoughts.
- Long-term stochastic behavior is naturally described by invariant measures and ergodic averages.
- Multispecies persistence is organized by average growth rates on boundary stochastic regimes.

Next reading: use Øksendal and Mao to consolidate the SDE tools, then move to research papers on stochastic Lotka–Volterra systems, epidemics, persistence, switching environments, or spatial stochastic models.