

Introduction to SDE Models in Population Biology

Lecture 0: Preliminaries for Beginners

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

Learning objectives and two-hour plan

By the end, students should be able to

- review the deterministic ODE ideas used in population models;
- use basic probability language: random variables, expectations, variances, independence;
- understand stochastic processes as random functions of time;
- read filtrations, adaptedness, stopping times, martingales, and Markov notation;
- understand the notation of an SDE before developing full Itô calculus.

Timing: 0–25 min ODE review; 25–55 min probability; 55–65 min break; 65–95 min stochastic processes; 95–115 min Brownian motion and SDE notation; 115–120 min summary.

Purpose of this preliminary lecture

The four main lectures use Brownian motion, Itô integrals, stochastic logistic equations, multispecies models, and invariant measures. This preliminary lecture slows down the beginning. It is intended for students who know differential equations but are beginners in stochastic processes.

Main purpose

Build common notation and intuition so that the later lectures can focus on SDE modeling, calculations, and qualitative dynamics.

Recommended background references

For a beginner-oriented short course, the following references are useful.

- L. J. S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*: friendly biological examples.
- B. Øksendal, *Stochastic Differential Equations*: standard concise introduction to Brownian motion and Itô calculus.
- X. Mao, *Stochastic Differential Equations and Applications*: strong solutions, stability, and Lyapunov methods.
- C. W. Gardiner, *Stochastic Methods*: useful intuition about diffusion and Fokker–Planck equations.

Deterministic scalar population models

A scalar population model often has the form

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0.$$

If $x(t) > 0$, then

$$\frac{x'(t)}{x(t)}$$

is the per-capita growth rate. **Example:**

$$x' = rx, \quad x(t) = x_0 e^{rt}.$$

Here r is the constant per-capita growth rate. Later, the SDE analogue will replace $r dt$ by a random increment.

Logistic growth: the basic nonlinear model

The logistic equation is

$$x' = x(a - bx), \quad a, b > 0.$$

Equilibria are

$$x = 0, \quad x^* = \frac{a}{b}.$$

For $x > 0$,

$$\frac{x'}{x} = a - bx.$$

Thus growth is positive below carrying capacity and negative above carrying capacity.

Why it matters

The stochastic logistic equation is the first nonlinear SDE model in the course.

Phase-line analysis: deterministic intuition

For

$$x' = x(a - bx),$$

the sign of $f(x) = x(a - bx)$ gives the dynamics:

$$0 < x < \frac{a}{b} : x' > 0, \quad x > \frac{a}{b} : x' < 0.$$

Therefore every positive solution approaches a/b . In SDEs, this phase-line picture is replaced by almost sure limits, time averages, invariant distributions, and stochastic persistence near the boundary.

Vector population models

An interacting population model often has the form

$$\frac{dx}{dt} = F(x), \quad x = (x_1, \dots, x_n) \in \mathbb{R}_+^n.$$

A two-species competition model is

$$\begin{cases} x' = x(a_1 - b_{11}x - b_{12}y), \\ y' = y(a_2 - b_{21}x - b_{22}y). \end{cases}$$

The positive cone $\mathbb{R}_+^n$ is biologically natural. In SDE models, preserving positivity is a major issue.

Equilibria and linearization

An equilibrium x^* satisfies

$$F(x^*) = 0.$$

Writing $z = x - x^*$ gives

$$z' = DF(x^*)z + \text{higher-order terms.}$$

In deterministic systems, eigenvalues of $DF(x^*)$ often determine local stability. In stochastic models, linearization is still important, but stability may mean stability in probability, almost sure stability, moment stability, or persistence in distribution.

Invariant regions and positivity

For population models, one usually wants

$$x_i(0) \geq 0 \implies x_i(t) \geq 0 \text{ for all } t \geq 0.$$

For the logistic ODE, positive solutions remain positive. For an SDE, positivity depends strongly on the noise term. Additive noise may push a population below zero, while multiplicative noise often respects the boundary:

$$dX_t = X_t(a - bX_t)dt + \sigma X_t dW_t.$$

Probability space: the sample-universe viewpoint

A probability space is written

$$(\Omega, \mathcal{F}, \mathbb{P}).$$

- Ω is the set of possible outcomes.
- $\mathcal{F}$ is the collection of events whose probabilities are defined.
- $\mathbb{P}$ assigns probabilities to events.

For SDEs, one outcome $\omega \in \Omega$ may represent one full environmental history. A solution is written as

$$X_t(\omega).$$

For fixed t , it is a random variable; for fixed ω , it is a sample path.

Events and sigma-algebras: beginner meaning

The symbol $\mathcal{F}$ is a sigma-algebra. For this course, the main interpretation is:

Interpretation

$\mathcal{F}$ records which questions about the random experiment are allowed and can be assigned probabilities.

If $A, B \in \mathcal{F}$, then

$$A^c, \quad A \cup B, \quad A \cap B$$

are also in $\mathcal{F}$. For stochastic processes, smaller sigma-algebras $\mathcal{F}_t$ represent information available up to time t .

Random variables

A real-valued random variable is a measurable function

$$X : \Omega \rightarrow \mathbb{R}.$$

Typical events are

$$\{X \leq a\}, \quad \{X > 0\}, \quad \{X \in I\}.$$

The distribution of X tells us the probabilities of such events.

Important point

A random variable is not just an unknown number. It is a function on the outcome space.

Distribution functions and densities

The distribution function of X is

$$F_X(x) = \mathbb{P}(X \leq x).$$

If X has density $p(x)$, then

$$F_X(x) = \int_{-\infty}^x p(y) dy, \quad \mathbb{P}(a < X < b) = \int_a^b p(y) dy.$$

For an SDE, the distribution of X_t changes with time. In many cases its density satisfies a Fokker–Planck equation.

Expectation

If X has density p , then

$$\mathbb{E}X = \int_{-\infty}^{\infty} xp(x) dx.$$

For a function g ,

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)p(x) dx.$$

In population biology, common quantities are

$$\mathbb{E}[X_t], \quad \mathbb{E}[\log X_t], \quad \mathbb{E}[X_t^2].$$

Usually

$$\mathbb{E}[f(X)] \neq f(\mathbb{E}X),$$

so nonlinearity matters.

Variance and covariance

The variance is

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2] - (\mathbb{E}X)^2.$$

It measures spread around the mean. The covariance of X and Y is

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)].$$

In multispecies models, correlated environmental noise may change coexistence and invasion thresholds.

Normal random variables

If $Z \sim N(0,1)$, then

$$\mathbb{E}Z = 0, \quad \text{Var}(Z) = 1.$$

If $X = \mu + \sigma Z$, then

$$X \sim N(\mu, \sigma^2).$$

Brownian increments are normal:

$$W_t - W_s \sim N(0, t - s).$$

Thus

$$W_{t+h} - W_t \sim \sqrt{h} Z.$$

This $\sqrt{h}$ scaling is the source of the new stochastic calculus rules.

Independence

Events A and B are independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Random variables are independent if information about one does not change the distribution of the other. Brownian motion has independent increments. For non-overlapping intervals,

$$W_{t_2} - W_{t_1} \quad \text{and} \quad W_{t_4} - W_{t_3}$$

are independent. This models fresh environmental shocks over time.

Conditional probability and conditional expectation

Conditional probability is

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \mathbb{P}(B) > 0.$$

Conditional expectation is a prediction using available information. In stochastic processes,

$$\mathbb{E}[X_t \mid \mathcal{F}_s], \quad s < t,$$

means the best prediction of the future random variable X_t using all information up to time s .

Law of large numbers and central limit idea

Let $Y_1, Y_2, \dots$ be independent identically distributed random variables with mean μ . The law of large numbers says

$$\frac{Y_1 + \dots + Y_n}{n} \rightarrow \mu$$

in a suitable probabilistic sense. The central limit idea says fluctuations around the mean are often approximately normal:

$$\frac{Y_1 + \dots + Y_n - n\mu}{\sqrt{n}} \approx \text{normal}.$$

Brownian motion can be viewed as a scaling limit of many small independent shocks.

Modes of convergence

For random variables X_n and X , several convergence notions occur.

- **Almost surely:** $X_n(\omega) \rightarrow X(\omega)$ for almost every ω .
- **In probability:** $\mathbb{P}(|X_n - X| > \varepsilon) \rightarrow 0$ for every $\varepsilon > 0$.
- **In mean square:** $\mathbb{E}|X_n - X|^2 \rightarrow 0$.
- **In distribution:** distributions converge in the weak sense.

Almost sure convergence is especially important for long-term population behavior.

Stochastic processes

A stochastic process is a family of random variables

$$X = (X_t)_{t \geq 0}.$$

For fixed t , X_t is a random variable. For fixed ω , the map

$$t \mapsto X_t(\omega)$$

is a sample path. For a population process, we ask:

$$\mathbb{P}(X_t > 0), \quad \mathbb{E}[X_t], \quad \lim_{t \rightarrow \infty} X_t, \quad \frac{1}{t} \int_0^t X_s ds.$$

Sample paths versus distributions

There are two complementary viewpoints.

- ① **Pathwise:** follow one realization $X_t(\omega)$.
- ② **Distributional:** study the probability law of X_t for fixed t .

For example, extinction may mean

$$X_t(\omega) \rightarrow 0 \quad \text{for almost every } \omega.$$

But one may also study whether the distribution of X_t approaches a density concentrated near zero.

Filtrations: information growing with time

A filtration is an increasing family of sigma-algebras

$$\mathcal{F}_s \subseteq \mathcal{F}_t, \quad 0 \leq s \leq t.$$

Interpretation:

$$\mathcal{F}_t = \text{information available up to time } t.$$

If W_t is Brownian motion, its natural filtration is generated by

$$W_s, \quad 0 \leq s \leq t.$$

A model should not use future noise to determine the present state.

Adapted processes

A process X_t is adapted to $(\mathcal{F}_t)$ if X_t is known from information available at time t . For SDEs, adaptedness means non-anticipation:

$$X_t \text{ may depend on } W_s, \quad 0 \leq s \leq t,$$

but not on future increments $W_{t+h} - W_t$. This is both mathematically essential for the Itô integral and biologically natural.

Stopping times

A random time τ is a stopping time if the event

$$\{\tau \leq t\}$$

can be decided using information up to time t . Example:

$$\tau_a = \inf\{t \geq 0 : X_t \leq a\}$$

is the first time a population falls below threshold a . Stopping times are used to localize estimates and to prove positivity or non-explosion.

Martingales

A process M_t is a martingale if

$$\mathbb{E}[|M_t|] < \infty, \quad \mathbb{E}[M_t \mid \mathcal{F}_s] = M_s, \quad 0 \leq s \leq t.$$

Informally, the best prediction of the future value is the current value. Brownian motion is a martingale:

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = W_s.$$

Stochastic integral terms often have martingale structure and mean zero.

Markov processes

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

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

Many SDE solutions are Markov processes. For population models, this means current population sizes contain all information needed for the probabilistic law of future population sizes.

From random walk to Brownian motion

Consider a simple random walk with many small independent increments. If time steps and space steps are scaled properly, the limiting process is Brownian motion. The key scaling is

$$\Delta W \approx \sqrt{\Delta t} Z, \quad Z \sim N(0, 1).$$

Biological intuition

Brownian motion represents accumulated environmental variation from many small independent shocks.

Brownian motion: definition for this course

A standard Brownian motion W_t satisfies:

- 1 $W_0 = 0$;
- 2 $W_t - W_s \sim N(0, t - s)$ for $0 \leq s < t$;
- 3 increments over disjoint intervals are independent;
- 4 sample paths are continuous almost surely.

Basic moments are

$$\mathbb{E}[W_t] = 0, \quad \mathbb{E}[W_t^2] = t.$$

For $s \leq t$,

$$\mathbb{E}[W_s W_t] = s.$$

Working with Brownian increments

For a small time step Δt ,

$$\Delta W = W_{t+\Delta t} - W_t$$

satisfies

$$\Delta W \sim N(0, \Delta t).$$

Thus in simulations,

$$\Delta W = \sqrt{\Delta t} Z, \quad Z \sim N(0, 1).$$

Useful moments are

$$\mathbb{E}[\Delta W] = 0, \quad \mathbb{E}[(\Delta W)^2] = \Delta t, \quad \mathbb{E}[(\Delta W)^4] = 3(\Delta t)^2.$$

Brownian paths are continuous but rough

Brownian paths are continuous, but not differentiable in the ordinary sense. Informally,

$$\frac{W_{t+h} - W_t}{h} \approx \frac{\sqrt{h}Z}{h} = \frac{Z}{\sqrt{h}},$$

which becomes large as $h \rightarrow 0$. Thus dW_t/dt is not an ordinary function. The notation dW_t must be interpreted through stochastic integration.

Quadratic variation

For a smooth deterministic function $g(t)$,

$$\sum_i (g(t_{i+1}) - g(t_i))^2 \rightarrow 0.$$

For Brownian motion,

$$\sum_i (W_{t_{i+1}} - W_{t_i})^2 \rightarrow t$$

in probability as the mesh goes to zero. This nonzero quadratic variation is the reason Itô's formula has an extra second-derivative term.

Differential notation for SDEs

The SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

means the integral equation

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s.$$

The first integral is an ordinary time integral. The second is an Itô stochastic integral. Remember: dt has size dt , while dW_t has typical size $\sqrt{dt}$.

Drift and diffusion coefficients

In

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

$b(x)$ is the **drift** and $\sigma(x)$ is the **diffusion coefficient**. For the stochastic logistic equation

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

we have

$$b(x) = x(a - bx), \quad \sigma(x) = \sigma x.$$

The noise amplitude is proportional to current population size.

Integral form of stochastic logistic growth

The SDE

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

means

$$X_t = X_0 + \int_0^t X_s(a - bX_s) ds + \int_0^t \sigma X_s dW_s.$$

The first integral represents accumulated deterministic density-dependent growth. The second integral represents accumulated environmental fluctuation. This model will be studied carefully in Lectures 1 and 2.

Existence and uniqueness: beginner statement

For

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

a standard theorem says: if b and σ are globally Lipschitz and have linear growth, then there is a unique global strong solution. Population models often have nonlinear coefficients that are not globally Lipschitz. Then one usually combines local existence with Lyapunov estimates to rule out explosion.

Positivity of population SDEs

Negative population sizes are biologically meaningless. A useful condition is that the diffusion coefficient vanishes at the boundary:

$$\sigma(0) = 0.$$

For multiplicative noise, $\sigma(x) = \sigma x$, so the noise amplitude is zero at $x = 0$. A common proof of positivity applies Itô's formula to $\log X_t$. If $\log X_t$ remains finite on finite intervals, then X_t cannot hit zero.

Itô formula: first version

For an ODE $dx = f(x)dt$, the ordinary chain rule gives

$$dV(x) = V'(x)f(x)dt.$$

For an SDE

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

Itô's formula gives

$$dV(X_t) = \left[V'(X_t)b(X_t) + \frac{1}{2}V''(X_t)\sigma^2(X_t) \right] dt + V'(X_t)\sigma(X_t)dW_t.$$

The second-derivative term is the main new feature.

Itô rule of thumb

A useful formal table is

$$(dt)^2 = 0, \quad (dt)(dW_t) = 0, \quad (dW_t)^2 = dt.$$

This is not ordinary algebra; it is a compact way to remember quadratic variation. If

$$dX_t = b dt + \sigma dW_t,$$

then

$$(dX_t)^2 \approx \sigma^2 (dW_t)^2 = \sigma^2 dt.$$

This produces the extra term in Itô's formula.

Worked example: logarithm transformation

Let

$$dX_t = X_t(a - bX_t)dt + \sigma X_t dW_t, \quad X_t > 0.$$

Set $V(x) = \log x$. Then

$$V'(x) = \frac{1}{x}, \quad V''(x) = -\frac{1}{x^2}.$$

Itô's formula gives

$$d \log X_t = \left(a - bX_t - \frac{\sigma^2}{2} \right) dt + \sigma dW_t.$$

Thus environmental noise lowers the effective logarithmic growth rate by $\sigma^2/2$.

Geometric Brownian motion

The SDE

$$dX_t = aX_t dt + \sigma X_t dW_t, \quad X_0 > 0,$$

is geometric Brownian motion. Using $V(x) = \log x$,

$$d \log X_t = \left(a - \frac{\sigma^2}{2} \right) dt + \sigma dW_t.$$

Therefore

$$X_t = X_0 \exp \left\{ \left(a - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\}.$$

This is the first explicit solution students should master.

Mean growth versus sample-path growth

For geometric Brownian motion,

$$\mathbb{E}[X_t] = X_0 e^{at}.$$

But along a typical sample path,

$$\frac{1}{t} \log X_t \rightarrow a - \frac{\sigma^2}{2} \quad \text{almost surely.}$$

Thus the mean may grow while typical sample paths decline if

$$a - \frac{\sigma^2}{2} < 0.$$

This is a key lesson: moment behavior and almost sure behavior can be very different.

Euler–Maruyama method

For

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

with time step Δt , the Euler–Maruyama scheme is

$$X_{n+1} = X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n,$$

where

$$\Delta W_n = \sqrt{\Delta t}Z_n, \quad Z_n \sim N(0, 1).$$

This method is useful for intuition, but it may fail to preserve positivity unless care is taken.

Choosing noise terms in biological models

A deterministic model $x' = xf(x)$ can be perturbed in several ways. **Environmental noise in per-capita growth:**

$$dX_t = X_t f(X_t) dt + \sigma X_t dW_t.$$

Demographic-type fluctuation:

$$dX_t = X_t f(X_t) dt + \sigma \sqrt{X_t} dW_t.$$

External additive forcing:

$$dX_t = X_t f(X_t) dt + \sigma dW_t.$$

The choice should reflect biological mechanism and boundary behavior.

Common beginner mistakes

- ① Treating dW_t/dt as an ordinary derivative.
- ② Forgetting the extra $\frac{1}{2} V'' \sigma^2$ term in Itô's formula.
- ③ Confusing $\mathbb{E}[X_t]$ with typical sample-path behavior.
- ④ Assuming positivity without checking the diffusion term near the boundary.
- ⑤ Using numerical methods without checking whether they preserve qualitative properties.
- ⑥ Thinking that stronger noise only changes variance; it may also change persistence thresholds.

In-class calculation 1

Let

$$dX_t = 2X_t dt + X_t dW_t, \quad X_0 = 1.$$

- 1 Find $d \log X_t$.
- 2 Write the explicit solution for X_t .
- 3 Compute the almost sure logarithmic growth rate.
- 4 Compute $\mathbb{E}[X_t]$.

Answer outline:

$$d \log X_t = \left(2 - \frac{1}{2}\right) dt + dW_t,$$

so

$$X_t = \exp \left\{ \frac{3}{2}t + W_t \right\}, \quad \mathbb{E}[X_t] = e^{2t}.$$

In-class calculation 2

Consider

$$dX_t = X_t(1 - X_t)dt + 0.8X_t dW_t.$$

Use Itô's formula with $V(x) = \log x$. **Calculation:**

$$d \log X_t = \left(1 - X_t - \frac{0.8^2}{2}\right) dt + 0.8dW_t.$$

Since $0.8^2/2 = 0.32$,

$$d \log X_t = (0.68 - X_t)dt + 0.8dW_t.$$

Near zero, the effective logarithmic growth rate is $0.68 > 0$.

Reading strategy for the main lectures

When reading a new SDE model, ask:

- ① What is the deterministic model when noise is removed?
- ② What biological mechanism does the noise represent?
- ③ Does the model preserve positivity?
- ④ What is the logarithmic growth rate near the boundary?
- ⑤ Is the long-time question about paths, moments, time averages, or invariant distributions?

This checklist will be used repeatedly in Lectures 1–4.

Summary and bridge to Lecture 1

In this preliminary lecture we reviewed:

- deterministic population-model language;
- random variables, expectation, variance, independence, and normal increments;
- stochastic processes, filtrations, stopping times, martingales, and Markov processes;
- Brownian motion and its $\sqrt{dt}$ scaling;
- SDE notation, drift, diffusion, positivity, and a first form of Itô's formula.

Lecture 1 begins Brownian motion, Itô integration, geometric Brownian motion, and the stochastic logistic equation.