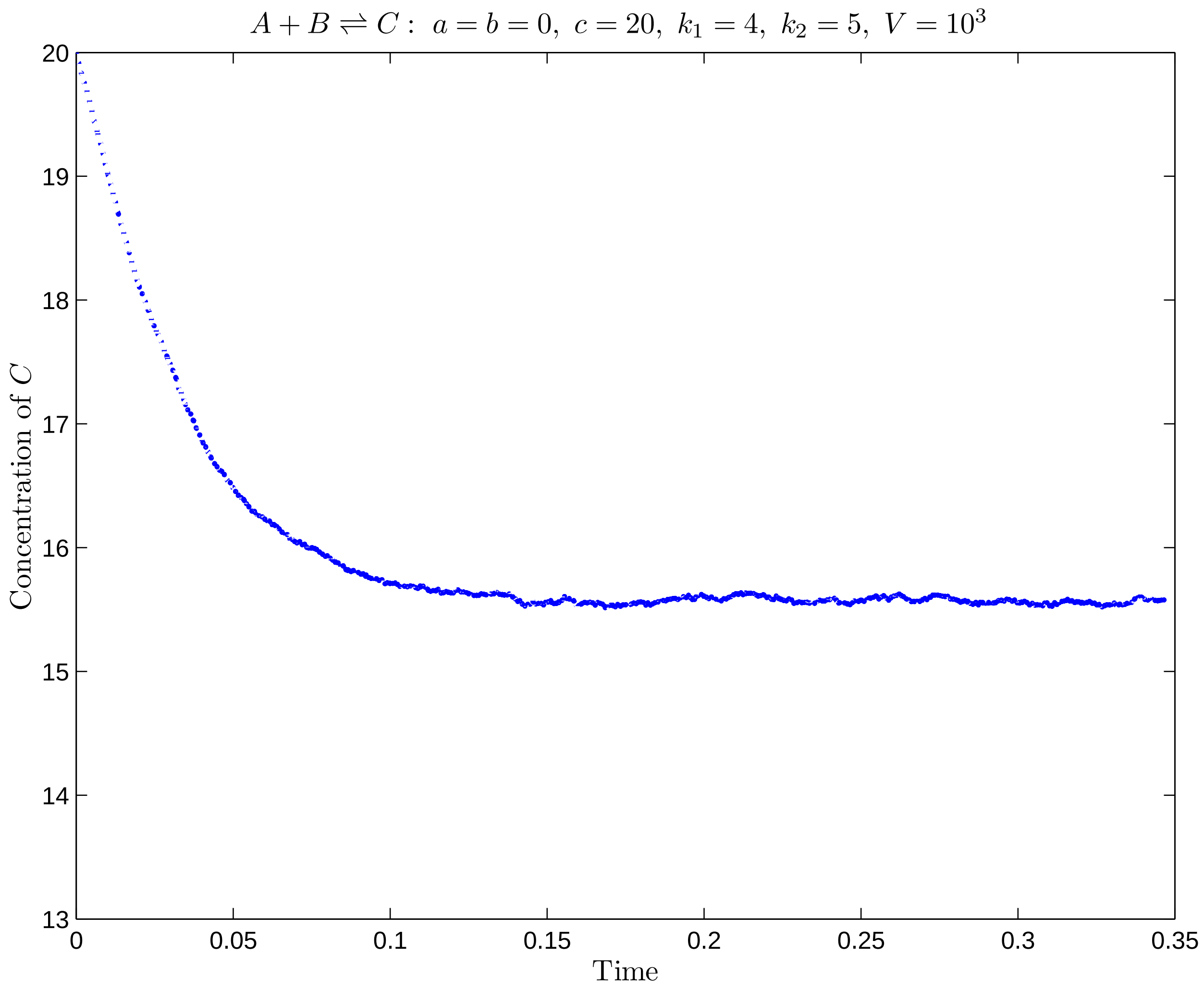
Stochastic models and their deterministic analogues

Phil Pollett

Department of Mathematics and MASCOS
University of Queensland

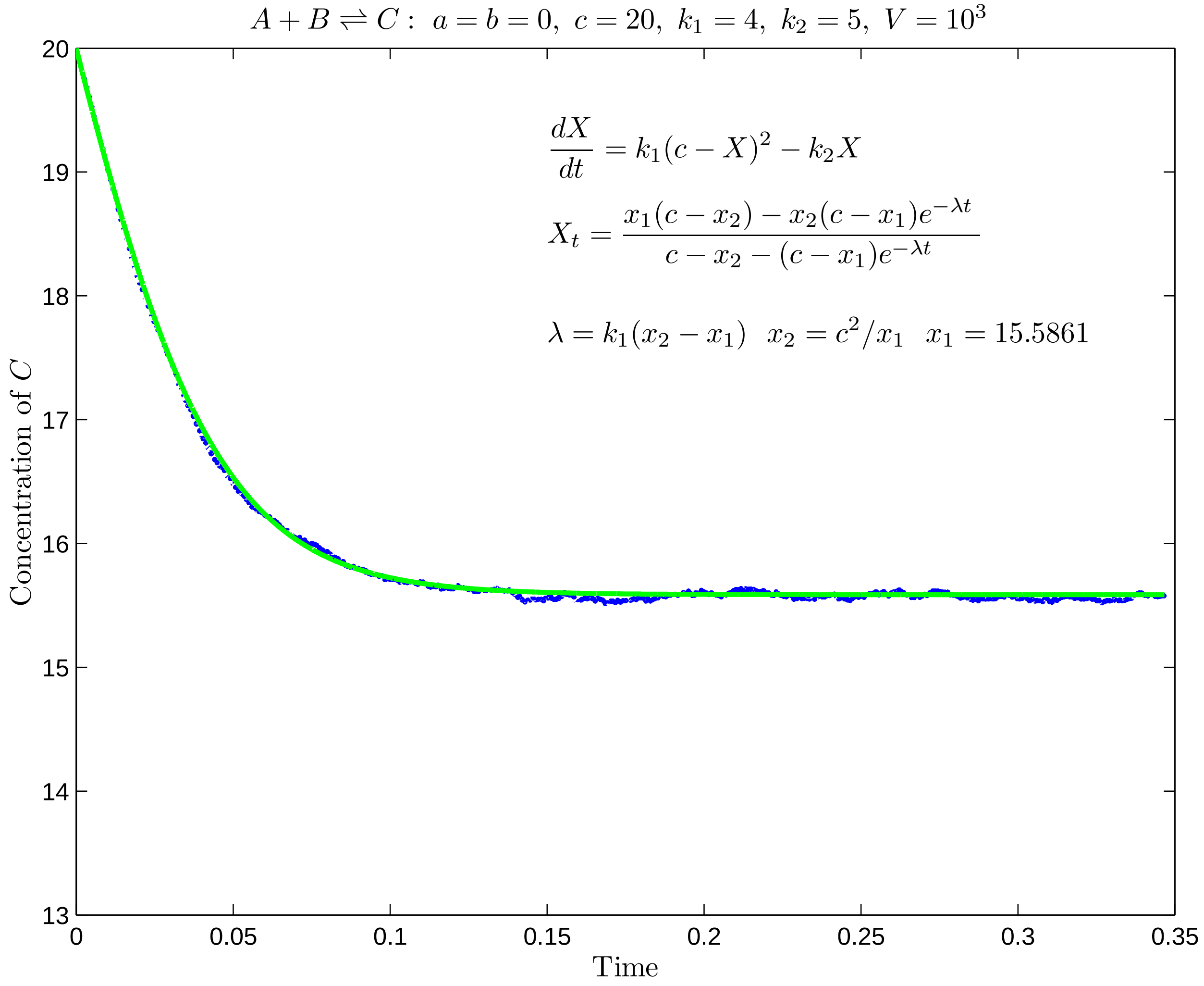
AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics
and Statistics of Complex Systems

A precipitation reaction



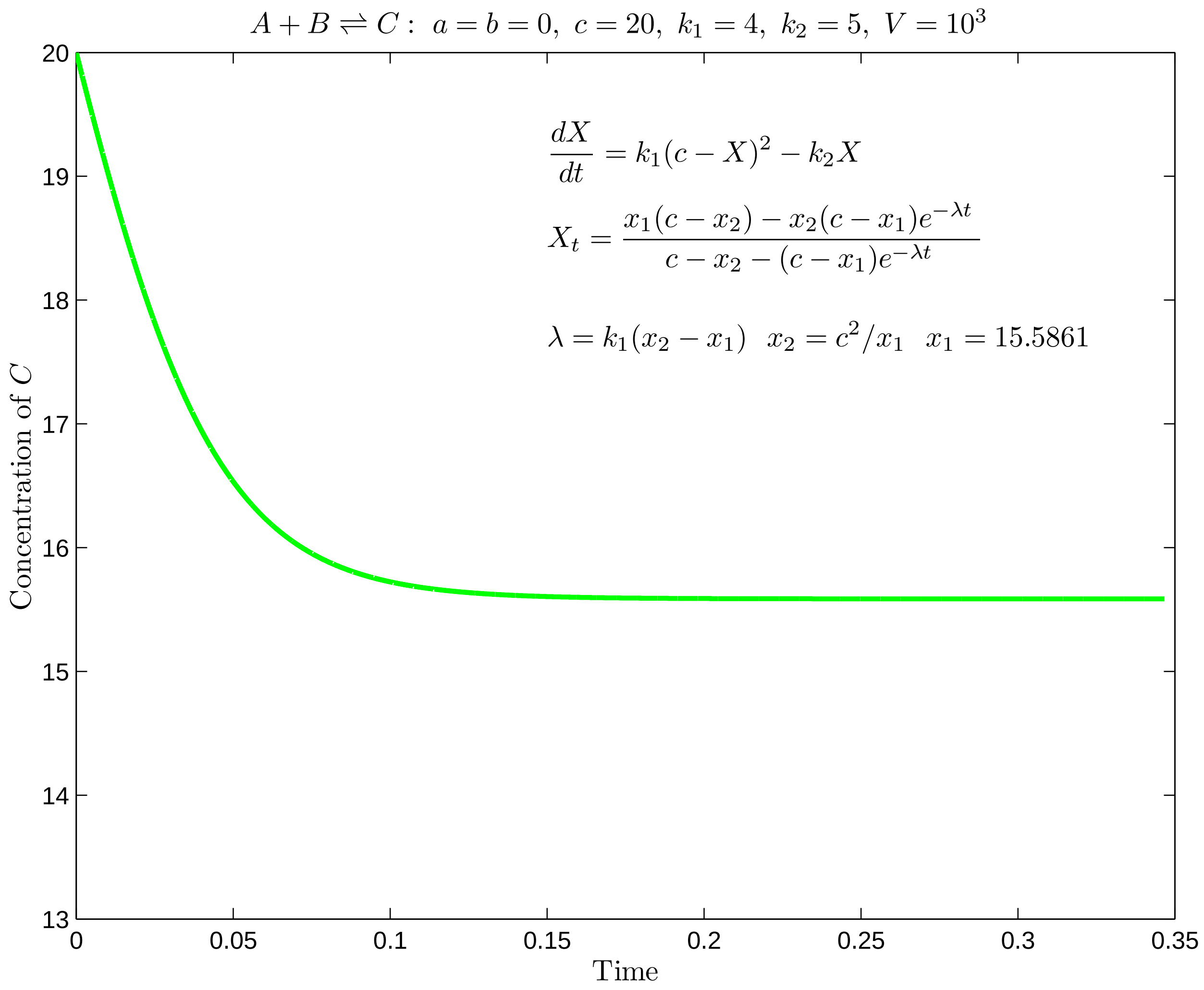
$$\text{Na}^+ + \text{Cl}^- \rightleftharpoons \text{NaCl}$$

A precipitation reaction



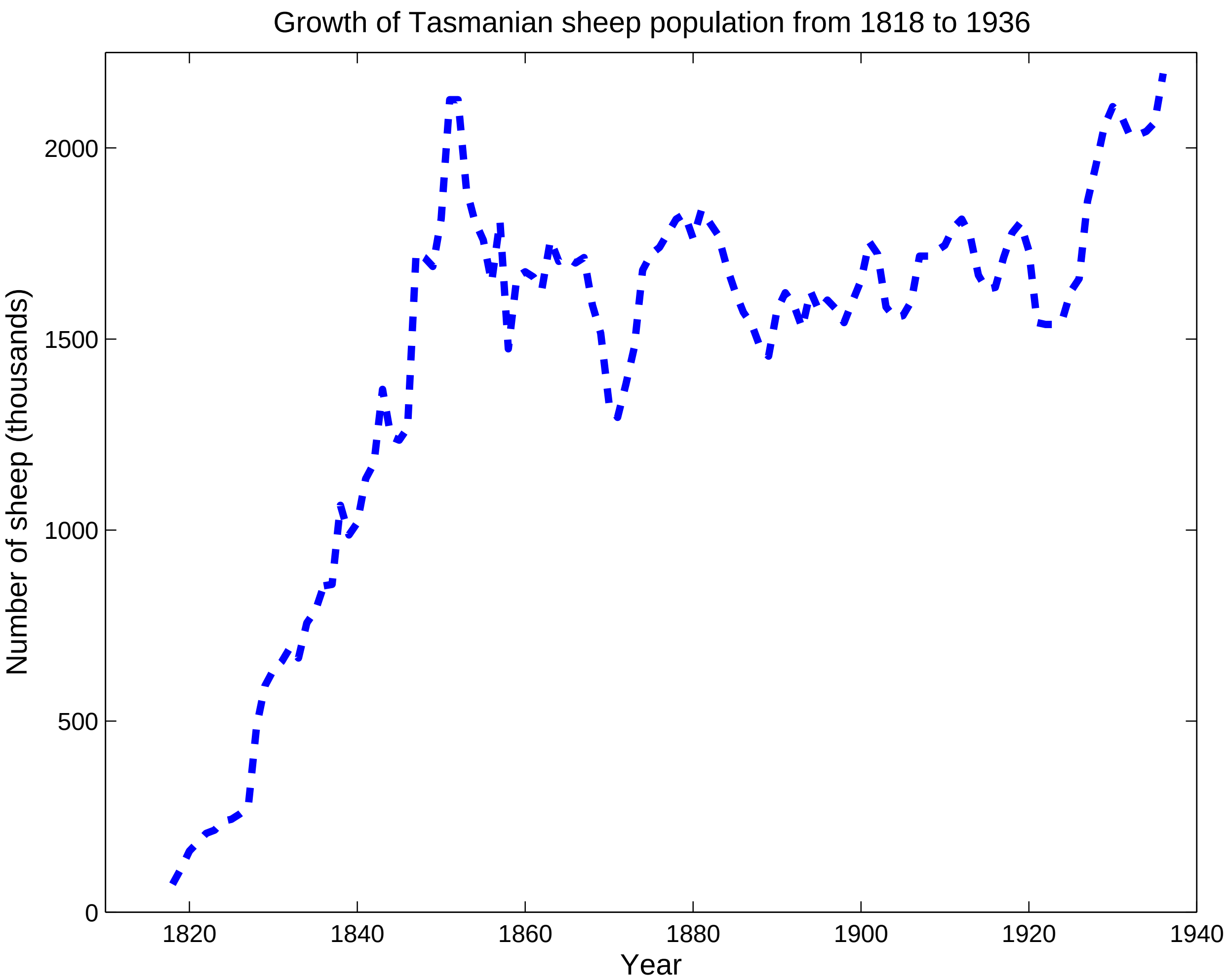
$Na^{+} + Cl^{-} \rightleftharpoons NaCl$

A precipitation reaction



$$Na^{+} + Cl^{-} \rightleftharpoons NaCl$$

Sheep in Tasmania



Davidson, J. (1938) On the growth of the sheep population in Tasmania, Trans. Roy. Soc. Sth. Austral. 62, 342–346.

A deterministic model

$$\frac{dn}{dt} = nf(n).$$

The net growth rate per individual is a function of the population size n.

We want $f(n)$ to be positive for small n and negative for large n.

A deterministic model

$$\frac{dn}{dt} = nf(n).$$

The net growth rate per individual is a function of the population size n.

We want $f(n)$ to be positive for small n and negative for large n. Simply set $f(n) = r - sn$ to give

$$\frac{dn}{dt} = n(r - sn).$$

A deterministic model

$$\frac{dn}{dt} = nf(n).$$

The net growth rate per individual is a function of the population size n.

We want $f(n)$ to be positive for small n and negative for large n. Simply set $f(n) = r - sn$ to give

$$\frac{dn}{dt} = n(r - sn).$$

This is the classical Verhulst* model (or *logistic model*):

*Verhulst, P.F. (1838) Notice sur la loi que la population suit dans son accroisement, *Corr. Math. et Phys.* X, 113–121.

The Verhulst model

Pierre Francois Verhulst (1804–1849, Brussels, Belgium)

The Verhulst model

Soit p la population : représentons par dp l'accroissement infiniment petit qu'elle reçoit pendant un temps infiniment court dt. Si la population croissait en progression géométrique, nous aurions l'équation $\frac{dp}{dt} = mp$. Mais comme la vitesse d'accroissement de la population est retardée par l'augmentation même du nombre des habitans, nous devrons retrancher de mp une fonction inconnue de p; de manière que la formule à intégrer deviendra

$$\frac{dp}{dt} = mp - \varphi(p).$$

L'hypothèse la plus simple que l'on puisse faire sur la forme de la fonction φ, est de supposer $\varphi(p) = np^2$. On trouve alors pour intégrale de l'équation ci-dessus

$$t = \frac{1}{m}\left[\log. p - \log.(m - np)\right] + \text{constante},$$

et il suffira de trois observations pour déterminer lés deux coefficiens constans m et n et la constante arbitraire.

The Verhulst model

En résolvant la dernière équation par rapport à p, il vient

$$p = \frac{mp' e^{mt}}{np' e^{mt} + m - np'} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

en désignant par p' la population qui répond à $t = o$, et par e la base des logarithmes népériens. Si l'on fait $t = \infty$, on voit que la valeur de p correspondante est $P = \frac{m}{n}$. Telle est donc *la limite supérieure de la population*.

Au lieu de supposer $\varphi p = np^2$, on peut prendre $\varphi p = np^\alpha$, α étant quelconque, ou $\varphi p = n \log. p$. Toutes ces hypothèses satisfont également bien aux faits observés ; mais elles donnent des valeurs très-différentes pour la limite supérieure de la population.

J'ai supposé successivement

$$\varphi p = np^2, \ \varphi p = np^3, \ \varphi p = np^4, \ \varphi p = n \log. p;$$

et les différences entre les populations calculées et celles que fournit l'observation ont été sensiblement les mêmes.

The Verhulst model

An alternative formulation has r being the growth rate with unlimited resources and K being the "natural" population size (the carrying capacity). We put $f(n) = r(1 - n/K)$ giving

$$\frac{dn}{dt} = rn(1 - n/K),$$

which is the original model with $s = r/K$.

The Verhulst model

An alternative formulation has r being the growth rate with unlimited resources and K being the “natural” population size (the carrying capacity). We put $f(n) = r(1 - n/K)$ giving

$$\frac{dn}{dt} = rn(1 - n/K),$$

which is the original model with $s = r/K$.

Integration gives

$$n_t = \frac{K}{1 + \left(\frac{K - n_0}{n_0}\right) e^{-rt}} \qquad (t \geq 0).$$

Verhulst-Pearl model

This formulation is due to Raymond Pearl:

Pearl, R. and Reed, L. (1920) On the rate of growth of population of the United States since 1790 and its mathematical representation, *Proc. Nat. Academy Sci.* 6, 275–288.

Pearl, R. (1925) *The biology of population growth*, Alfred A. Knopf, New York.

Pearl, R. (1927) The growth of populations, *Quart. Rev. Biol.* 2, 532–548.

Verhulst-Pearl model

Raymond Pearl (1879–1940, Farmington, N.H., USA)

Pearl was a “social drinker”

Pearl was widely known for his lust for life and his love of food, drink, music and parties. He was a key member of the Saturday Night Club. Prohibition made no dent in Pearl’s drinking habits (which were legendary).

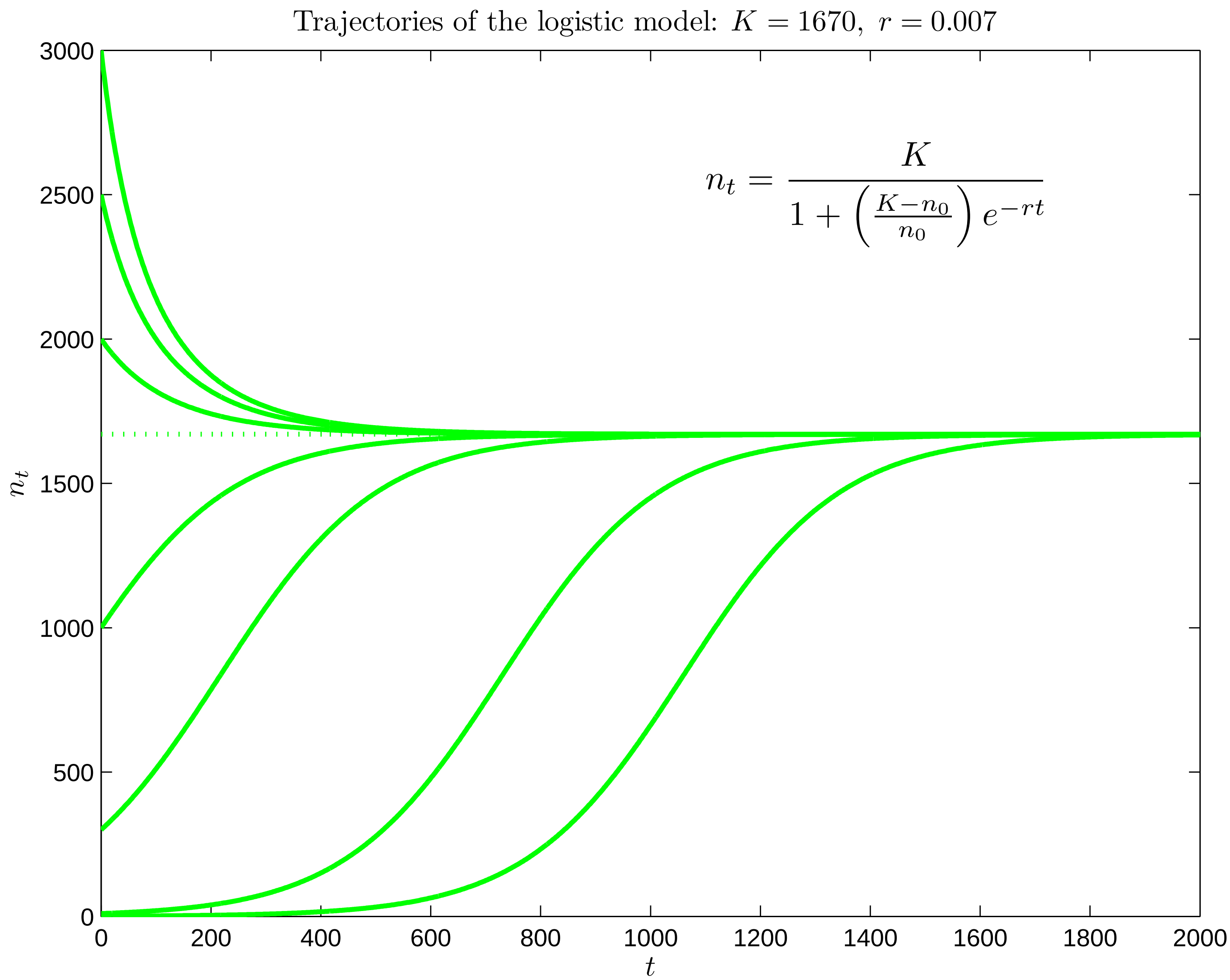
Pearl was a "social drinker"

Pearl was widely known for his lust for life and his love of food, drink, music and parties. He was a key member of the Saturday Night Club. Prohibition made no dent in Pearl's drinking habits (which were legendary).

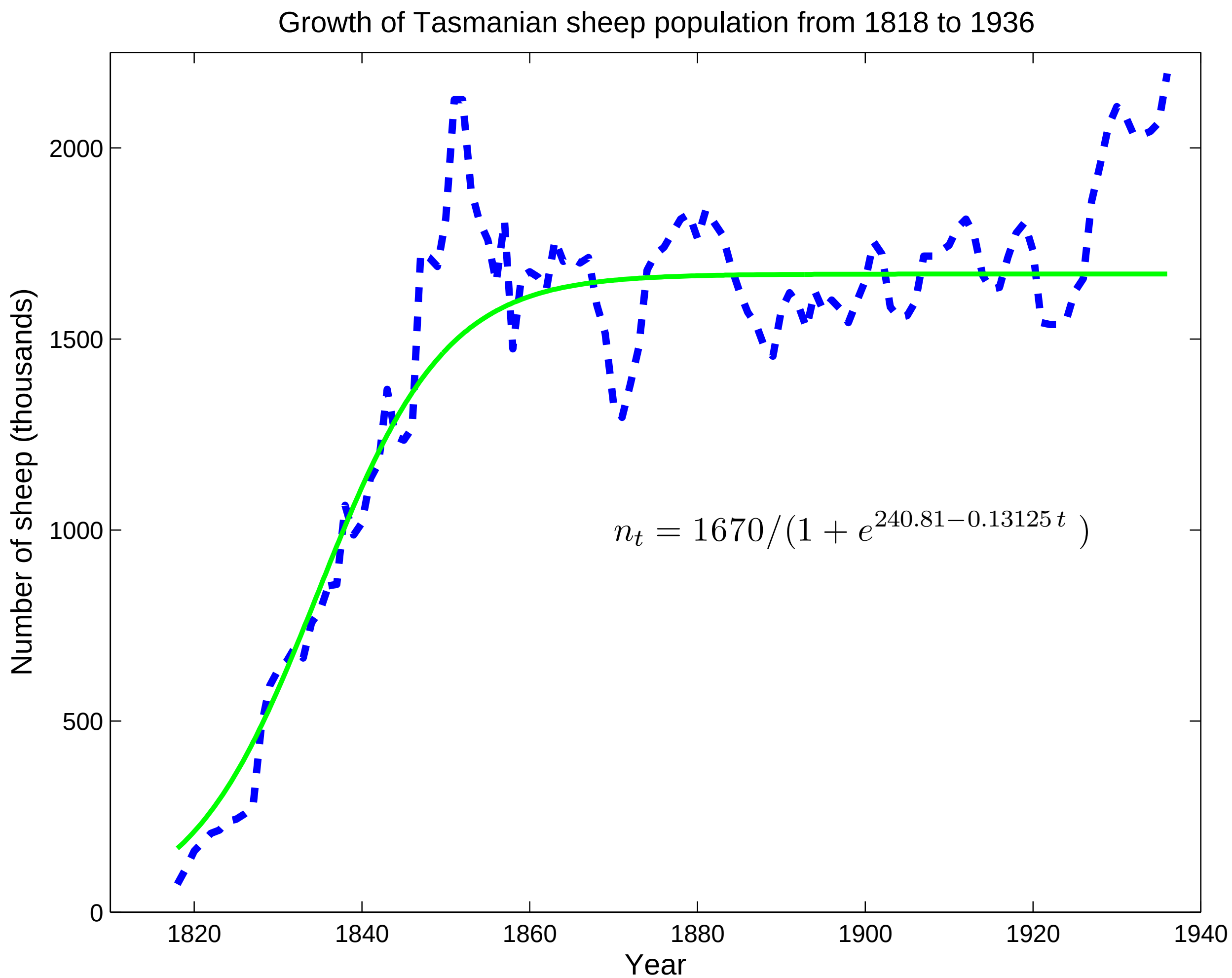
In 1926, his book, Alcohol and Longevity, demonstrated that drinking alcohol in moderation is associated with greater longevity than either abstaining or drinking heavily.

Pearl, R. (1926) *Alcohol and Longevity*, Alfred A. Knopf, New York.

Verhulst-Pearl model

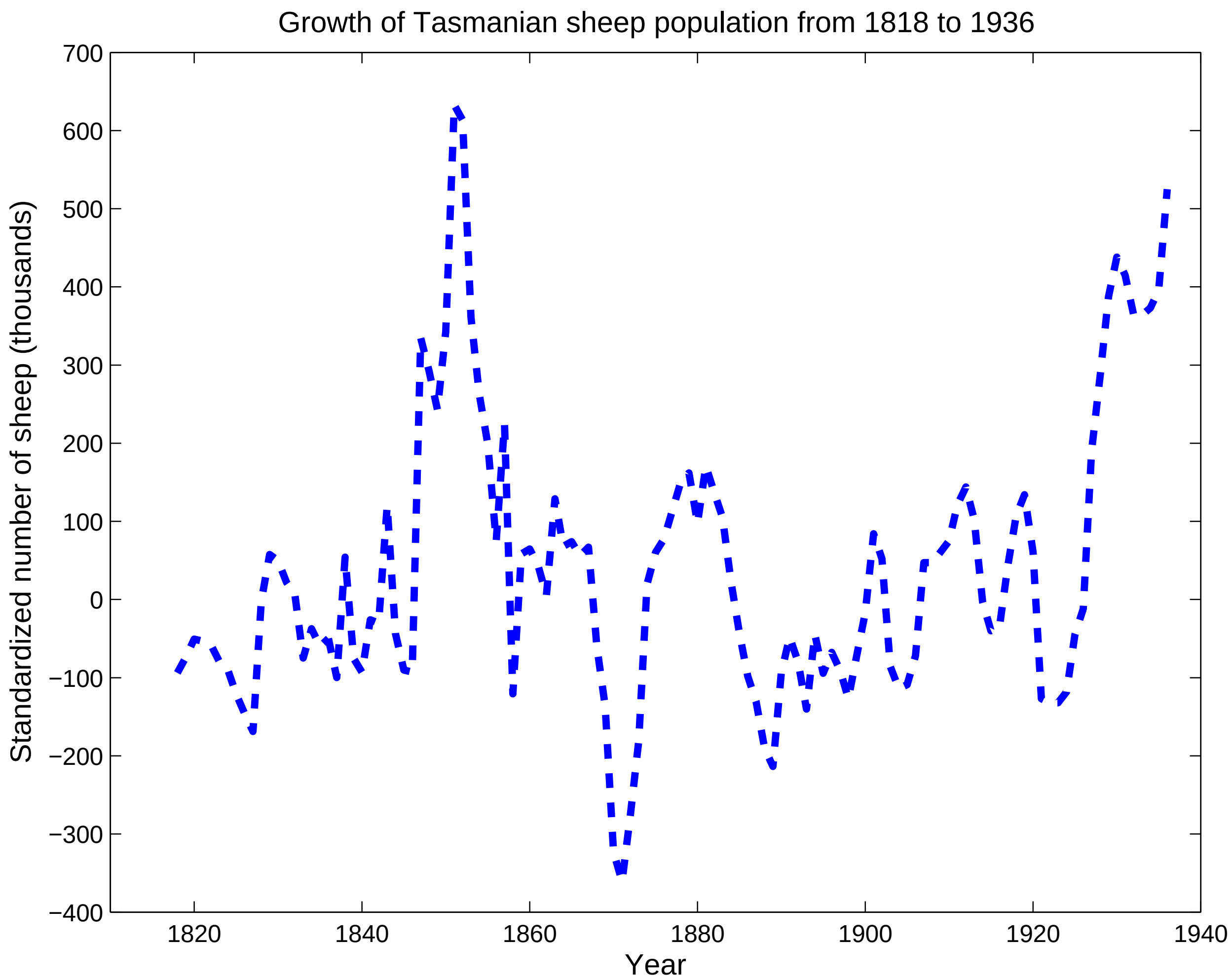


Sheep in Tasmania



Davidson, J. (1938) On the growth of the sheep population in Tasmania, Trans. Roy. Soc. Sth. Austral. 62, 342–346.

Sheep in Tasmania



(With the deterministic trajectory subtracted)

A stochastic model

We really need to account for the variation observed.

A recent approach to stochastic modelling in Applied Mathematics can be summarised as follows:

> “I feel guilty – I should add some noise”

(promulgated by stochastic modelling “experts” and courses in Financial Mathematics that require no background in stochastic processes).

A stochastic model

We really need to account for the variation observed.

A recent approach to stochastic modelling in Applied Mathematics can be summarised as follows:

> "I feel guilty – I should add some noise"

(promulgated by stochastic modelling "experts" and courses in Financial Mathematics that require no background in stochastic processes)*.

*Zen Maxim (for survival in a modern university): Before you criticize someone, you should walk a mile in their shoes. That way, when you criticize them, you're a mile away and you have their shoes.

Adding noise

In our case,

$$n_t = \frac{K}{1 + \left(\frac{K - n_0}{n_0}\right) e^{-rt}} + \text{something random}$$

or perhaps

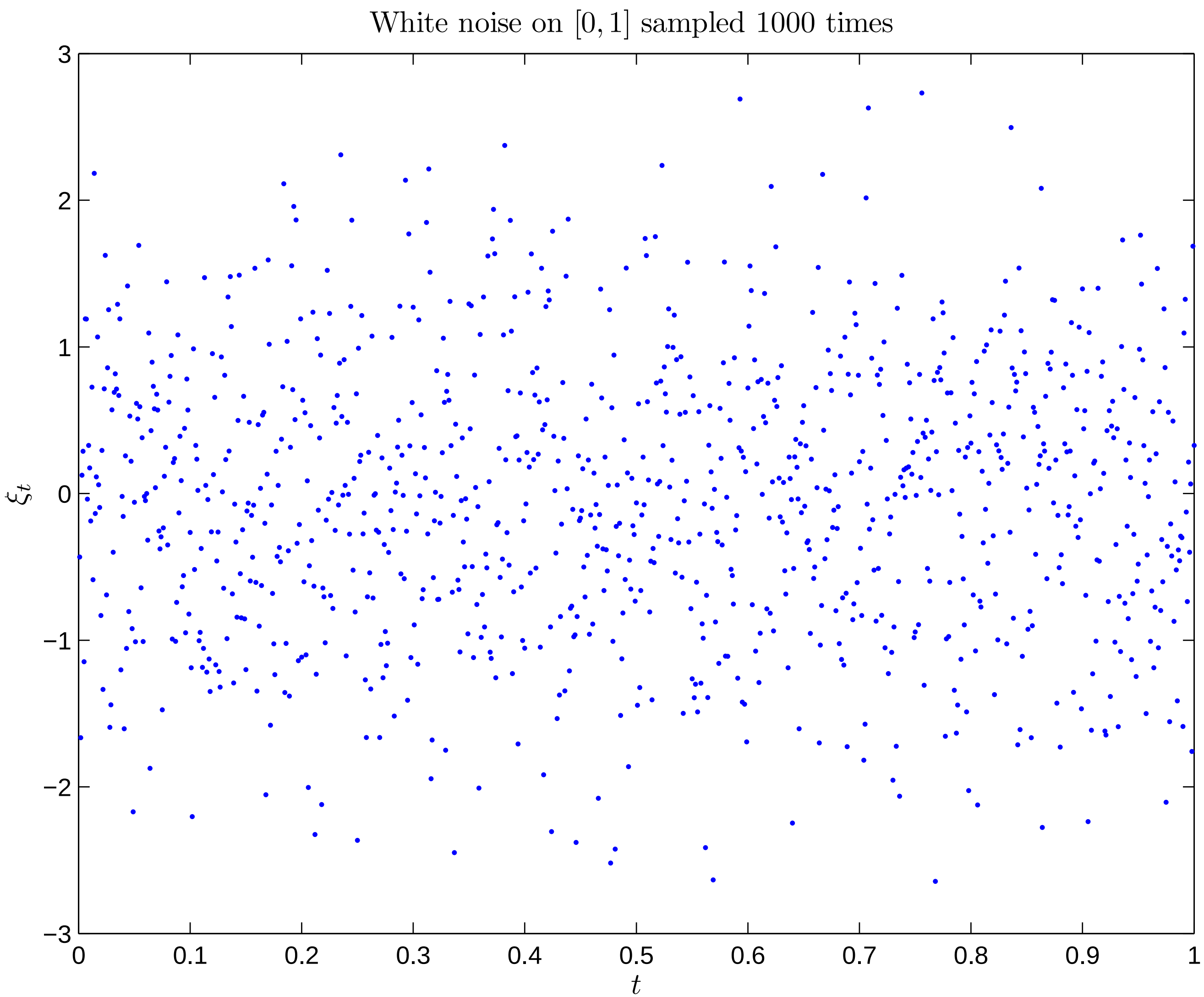
$$\frac{dn}{dt} = rn\left(1 - \frac{n}{K}\right) + \sigma \times \text{noise.}$$

Noise?

The usual model for “noise” is *white noise* (or *pure Gaussian noise*).

Imagine a random process $(\xi_t,\, t \geq 0)$ with $\xi_t \sim N(0,1)$ for all t and $\xi_{t_1}, \ldots, \xi_{t_n}$ *independent* for all finite sequences of times $t_1, \ldots, t_n$.

White noise



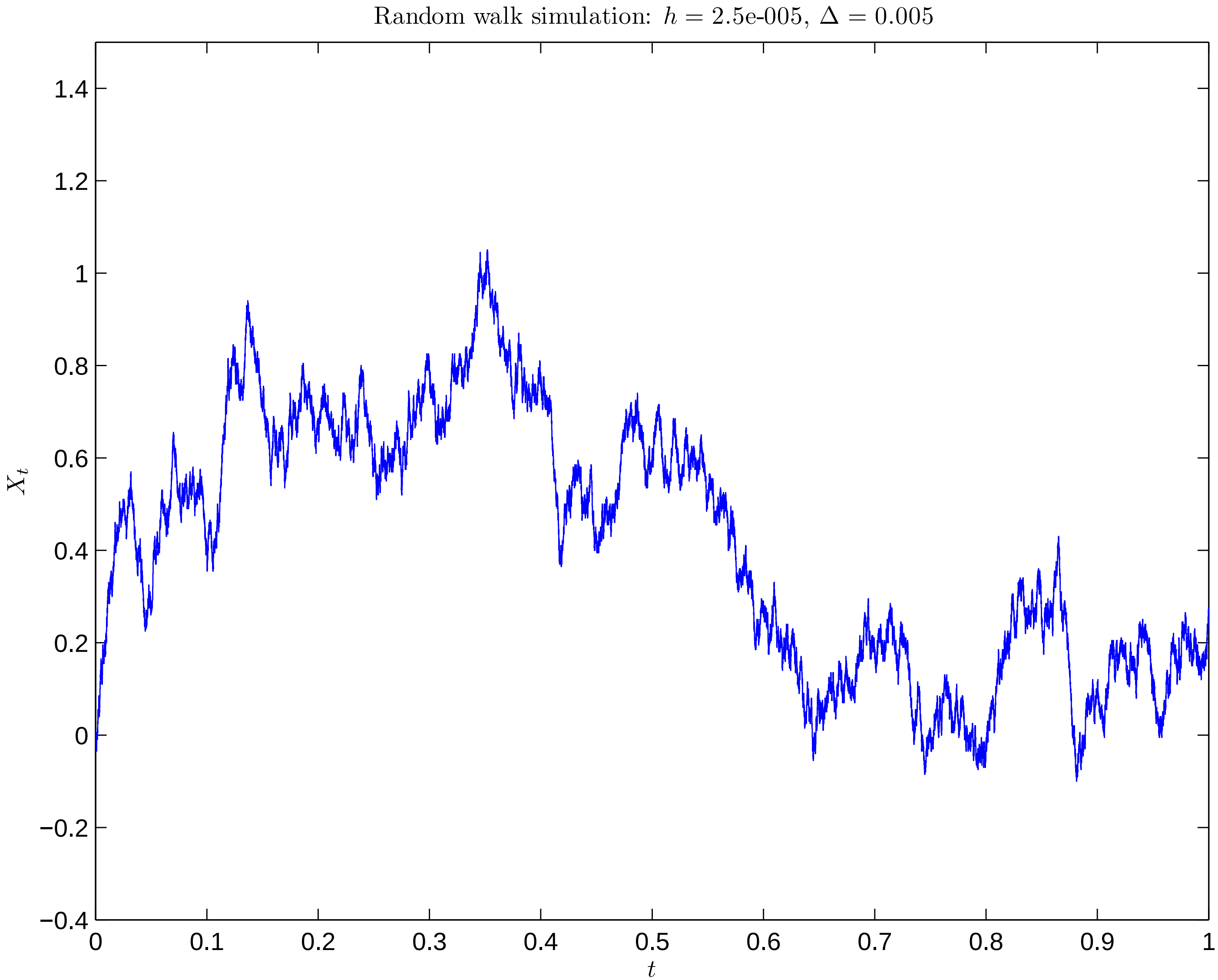
White noise on $[0,1]$ sampled 1000 times
3
2
1
0
−1
−2
−3
ξ_t
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
t

Brownian motion

The white noise process $(\xi_t,\, t \geq 0)$ is formally defined as the derivative of *standard Brownian motion* $(B_t,\, t \geq 0)$.

Brownian motion (or Wiener process) can be constructed by way of a random walk. A particle starts at 0 and takes small steps of size $+\Delta$ or $-\Delta$ with equal probability $p = 1/2$ after successive time steps of size h. If $\Delta \sim \sqrt{h}$, as $h \to 0$, then the limit process is *standard Brownian motion*.

Symmetric random walk: $\Delta = \sqrt{h}$



Brownian motion

The white noise process $(\xi_t,\, t \geq 0)$ is formally defined as the derivative of *standard Brownian motion* $(B_t,\, t \geq 0)$.

Brownian motion (or Wiener process) can be constructed by way of a random walk. A particle starts at 0 and takes small steps of size $+\Delta$ or $-\Delta$ with equal probability $p = 1/2$ after successive time steps of size h. If $\Delta \sim \sqrt{h}$, as $h \to 0$, then the limit process is *standard Brownian motion*.

Brownian motion

The white noise process $(\xi_t,\, t \geq 0)$ is formally defined as the derivative of *standard Brownian motion* $(B_t,\, t \geq 0)$.

Brownian motion (or Wiener process) can be constructed by way of a random walk. A particle starts at 0 and takes small steps of size $+\Delta$ or $-\Delta$ with equal probability $p = 1/2$ after successive time steps of size h. If $\Delta \sim \sqrt{h}$, as $h \to 0$, then the limit process is *standard Brownian motion*.

This construction permits us to write $dB_t = \xi_t \sqrt{dt}$, with the interpretation that a change in B_t in time dt is a Gaussian random variable with $\mathbb{E}(dB_t) = 0$, $\mathrm{Var}(dB_t) = dt$ and $\mathrm{Cov}(dB_t, dB_s) = 0\ (s \neq t)$.

Brownian motion

The white noise process $(\xi_t,\, t \geq 0)$ is formally defined as the derivative of *standard Brownian motion* $(B_t,\, t \geq 0)$.

Brownian motion (or Wiener process) can be constructed by way of a random walk. A particle starts at 0 and takes small steps of size $+\Delta$ or $-\Delta$ with equal probability $p = 1/2$ after successive time steps of size h. If $\Delta \sim \sqrt{h}$, as $h \to 0$, then the limit process is *standard Brownian motion*.

This construction permits us to write $dB_t = \xi_t \sqrt{dt}$, with the interpretation that a change in B_t in time dt is a Gaussian random variable with $\mathbb{E}(dB_t) = 0$, $\mathrm{Var}(dB_t) = dt$ and $\mathrm{Cov}(dB_t, dB_s) = 0\ (s \neq t)$.

The correct interpretation is by way of the Itô integral:

$$B_t = \int_0^t dB_s = \int_0^t \xi_s\, ds.$$

Brownian motion

General Brownian motion $(W_t,\, t \geq 0)$, with drift μ and variance σ^2, can be constructed in the same way but with $\Delta \sim \sigma\sqrt{h}$ and $p = \frac{1}{2}\left(1 + (\mu/\sigma)\sqrt{h}\right)$, and we may write

$$dW_t = \mu\, dt + \sigma\, dB_t,$$

with the interpretation that a change in W_t in time dt is a Gaussian random variable with $\mathbb{E}(dW_t) = \mu dt$, $\mathrm{Var}(dW_t) = \sigma^2 dt$ and $\mathrm{Cov}(dW_t, dW_s) = 0$.

Brownian motion

General Brownian motion $(W_t,\, t \geq 0)$, with drift μ and variance σ^2, can be constructed in the same way but with $\Delta \sim \sigma\sqrt{h}$ and $p = \frac{1}{2}\left(1 + (\mu/\sigma)\sqrt{h}\right)$, and we may write

$$dW_t = \mu\, dt + \sigma\, dB_t,$$

with the interpretation that a change in W_t in time dt is a Gaussian random variable with $\mathbb{E}(dW_t) = \mu dt$, $\mathrm{Var}(dW_t) = \sigma^2 dt$ and $\mathrm{Cov}(dW_t, dW_s) = 0$. This *stochastic differential equation (SDE)* can be integrated to give $W_t = \mu t + \sigma B_t$.

Brownian motion

General Brownian motion $(W_t,\, t \geq 0)$, with drift μ and variance σ^2, can be constructed in the same way but with $\Delta \sim \sigma\sqrt{h}$ and $p = \frac{1}{2}\left(1 + (\mu/\sigma)\sqrt{h}\right)$, and we may write

$$dW_t = \mu\, dt + \sigma\, dB_t,$$

with the interpretation that a change in W_t in time dt is a Gaussian random variable with $\mathbb{E}(dW_t) = \mu dt$, $\mathrm{Var}(dW_t) = \sigma^2 dt$ and $\mathrm{Cov}(dW_t, dW_s) = 0$. This *stochastic differential equation (SDE)* can be integrated to give $W_t = \mu t + \sigma B_t$.

It does not require an enormous leap of faith for us now to write down, and properly interpret, the SDE

$$dn_t = r n_t \left(1 - n_t/K\right) dt + \sigma dB_t$$

as a model for growth of our sheep population.

Adding noise

The idea (indeed the very idea of an SDE) can be traced back to Paul Langevin's 1908 paper "On the theory of Brownian Motion":

> Langevin, P. (1908) Sur la théorie du mouvement brownien, *Comptes Rendus* 146, 530–533.

He derived a "dynamic theory" of Brownian Motion three years after Einstein's ground breaking paper on Brownian Motion:

> Einstein, A. (1905) On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat, *Ann. Phys.* 17, 549–560 [English translation by Anna Beck in *The Collected Papers of Albert Einstein*, Princeton University Press, Princeton, USA, 1989, Vol. 2, pp. 123–134.]

Langevin

Langevin introduced a "stochastic force" (his phrase "complementary force"–complimenting the viscous drag μ) pushing the Brownian particle around in velocity space (Einstein worked in configuration space).

In modern terminology, Langevin described the Brownian particle's velocity as an *Ornstein-Uhlenbeck (OU) process* and its position as the time integral of its velocity, while Einstein described its position as a Wiener process.

The *Langevin equation* (for a particle of unit mass) is

$$dv_t = -\mu v_t \, dt + \sigma dB_t.$$

This is Newton's law ($-\mu v =$ Force $= m\dot{v}$) *plus* noise. The (strong) solution to this SDE is the OU process.

Langevin

Langevin introduced a “stochastic force” (his phrase “complementary force”–complimenting the viscous drag μ) pushing the Brownian particle around in velocity space (Einstein worked in configuration space).

In modern terminology, Langevin described the Brownian particle’s velocity as an *Ornstein-Uhlenbeck (OU) process* and its position as the time integral of its velocity, while Einstein described its position as a Wiener process.

The *Langevin equation* (for a particle of unit mass) is

$$dv_t = -\mu v_t \, dt + \sigma dB_t.$$

This is Newton’s law ($-\mu v =$ Force $= m\dot{v}$) *plus* noise. The (strong) solution to this SDE is the OU process. Warning: $\int_0^t v_s \, ds \neq B_t$; this functional is not even Markovian.

Langevin

Einstein said of Langevin "... It seems to me certain that he would have developed the special theory of relativity if that had not been done elsewhere, for he had clearly recognized the essential points."

Paul Langevin (1872-1946, Paris, France)

Langevin was a dark horse

In 1910 he had an affair with *Marie Curie* (Polish physicist).

Langevin was a dark horse

In 1910 he had an affair with *Marie Curie* (Polish physicist).

Langevin was a dark horse

In 1910 he had an affair with *Marie Curie* (Polish physicist).

The person on the right is not Langevin, but Langevin's PhD supervisor *Pierre Curie*.

Solution to Langevin's equation

To solve $dv_t = -\mu v_t \, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$.

Solution to Langevin's equation

To solve $dv_t = -\mu v_t\, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$.
Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

Solution to Langevin's equation

To solve $dv_t = -\mu v_t \, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$. Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

But, from Langevin's equation we have that

$$e^{\mu t} dv_t = -\mu e^{\mu t} v_t \, dt + \sigma e^{\mu t} dB_t,$$

Solution to Langevin's equation

To solve $dv_t = -\mu v_t \, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$. Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

But, from Langevin's equation we have that

$$e^{\mu t} dv_t = -\mu e^{\mu t} v_t \, dt + \sigma e^{\mu t} dB_t,$$

and hence that $dy_t = \sigma e^{\mu t} dB_t$.

Solution to Langevin's equation

To solve $dv_t = -\mu v_t\, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$. Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

But, from Langevin's equation we have that

$$e^{\mu t} dv_t = -\mu e^{\mu t} v_t\, dt + \sigma e^{\mu t} dB_t,$$

and hence that $dy_t = \sigma e^{\mu t} dB_t$. Integration gives

$$y_t = y_0 + \int_0^t \sigma e^{\mu s} dB_s,$$

Solution to Langevin's equation

To solve $dv_t = -\mu v_t \, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$. Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

But, from Langevin's equation we have that

$$e^{\mu t} dv_t = -\mu e^{\mu t} v_t \, dt + \sigma e^{\mu t} dB_t,$$

and hence that $dy_t = \sigma e^{\mu t} dB_t$. Integration gives

$$y_t = y_0 + \int_0^t \sigma e^{\mu s} dB_s,$$

and so (the *Ornstein-Uhlenbeck process*)

$$v_t = v_0 e^{-\mu t} + \int_0^t \sigma e^{-\mu(t-s)} dB_s.$$

Solution to Langevin's equation

To solve $dv_t = -\mu v_t\, dt + \sigma dB_t$, consider the process $y_t = v_t e^{\mu t}$. Differentiation (Itô calculus!) gives $dy_t = e^{\mu t} dv_t + \mu e^{\mu t} v_t dt$.

But, from Langevin's equation we have that

$$e^{\mu t} dv_t = -\mu e^{\mu t} v_t\, dt + \sigma e^{\mu t} dB_t,$$

and hence that $dy_t = \sigma e^{\mu t} dB_t$. Integration gives

$$y_t = y_0 + \int_0^t \sigma e^{\mu s} dB_s,$$

and so (the *Ornstein-Uhlenbeck process*)

$$v_t = v_0 e^{-\mu t} + \int_0^t \sigma e^{-\mu(t-s)} dB_s.$$

We can deduce much from this. For example, v_t is a Gaussian process with $\mathbb{E}(v_t) = v_0 e^{-\mu t}$ and $\mathrm{Var}(v_t) = \frac{\sigma^2}{2\mu}(1 - e^{-2\mu t})$, and

$$\mathrm{Cov}(v_t, v_{t+s}) = \mathrm{Var}(v_t) e^{-\mu|s|}.$$

Where were we?

We had just added noise to our logistic model:

$$dn_t = rn_t\left(1 - \frac{n_t}{K}\right)dt + \sigma\, dB_t. \tag{1}$$

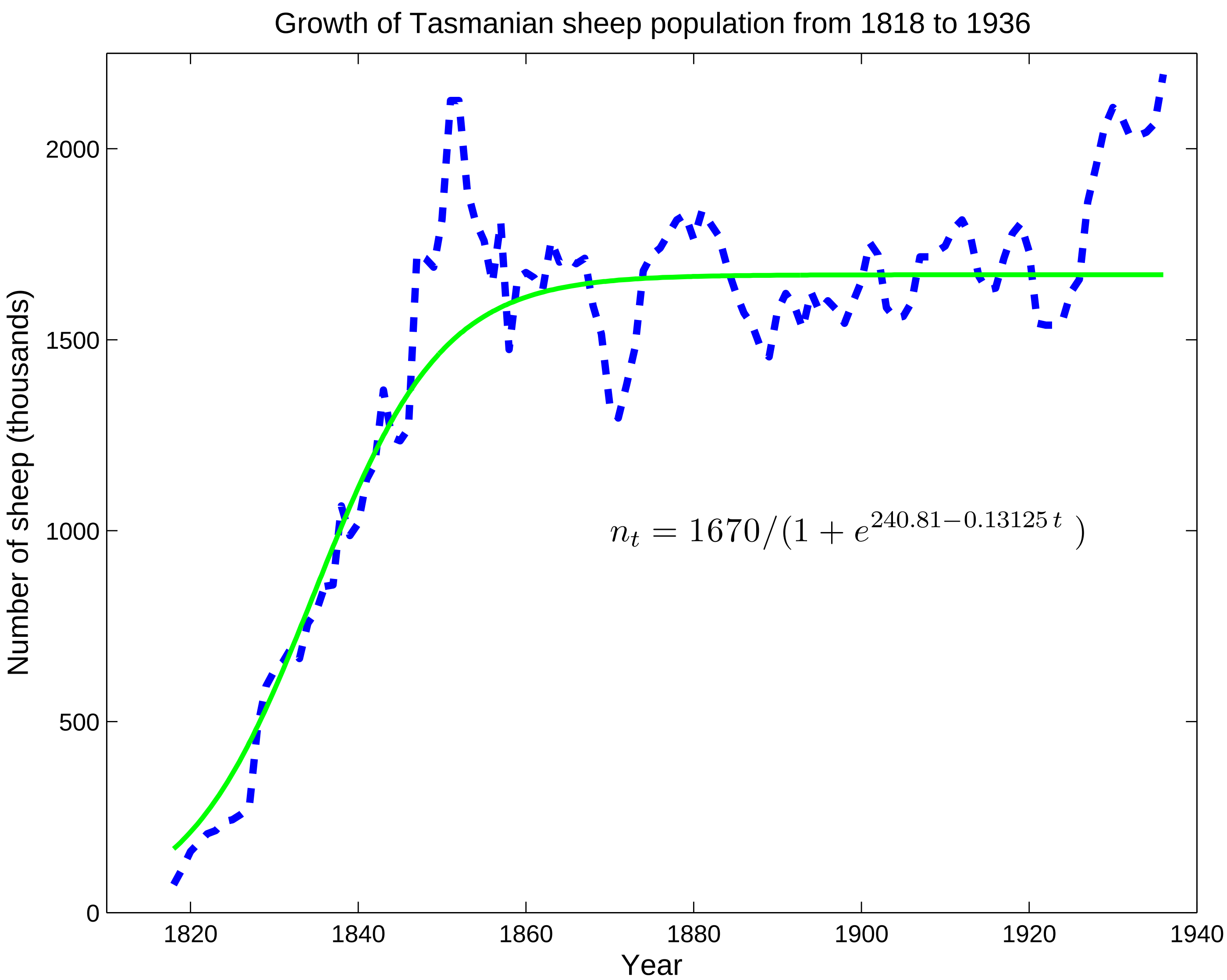
Where were we?

We had just added noise to our logistic model:

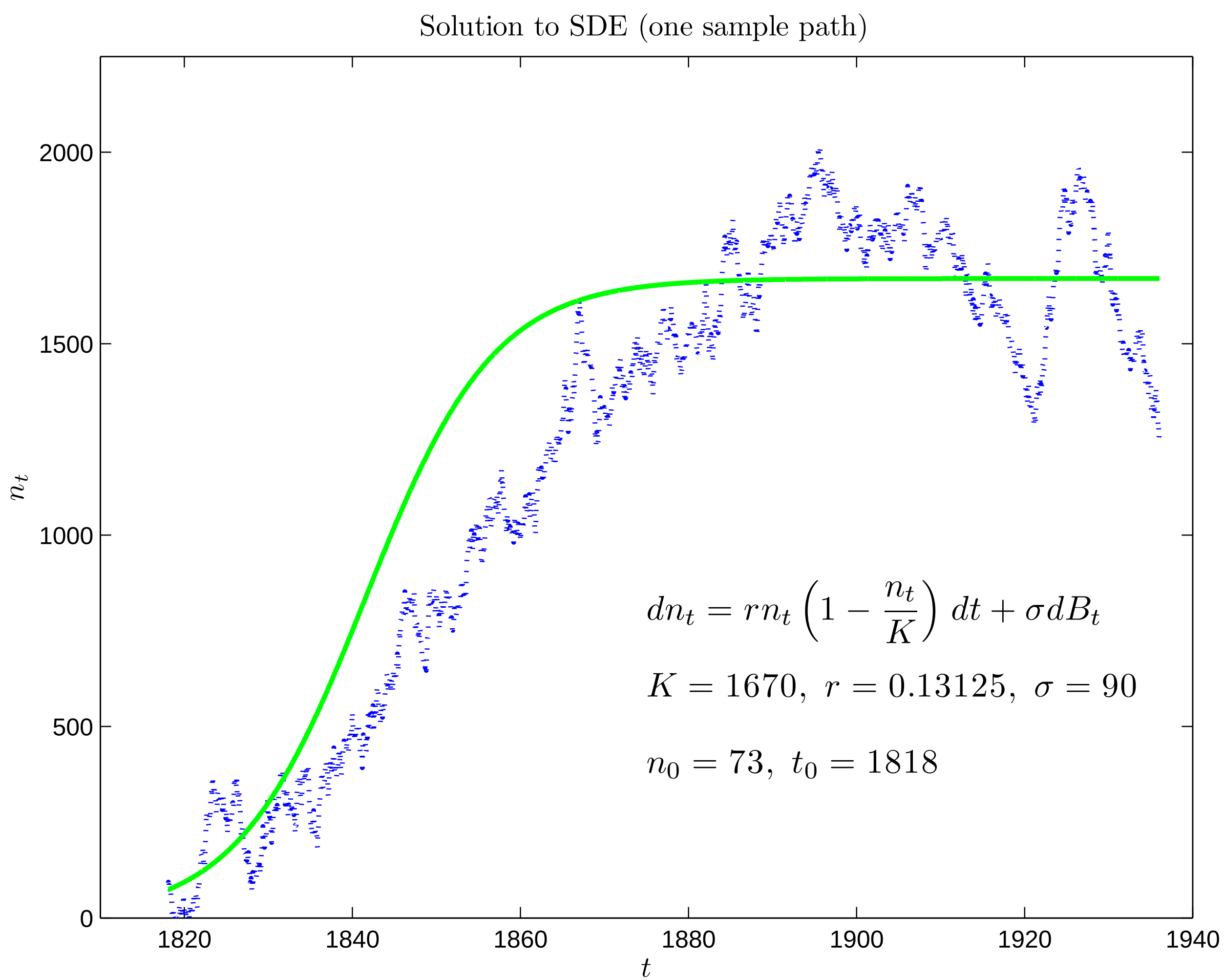
$$dn_t = rn_t \left(1 - \frac{n_t}{K}\right) dt + \sigma \, dB_t. \tag{1}$$

So, what the hell is wrong with (1)?

Sheep in Tasmania

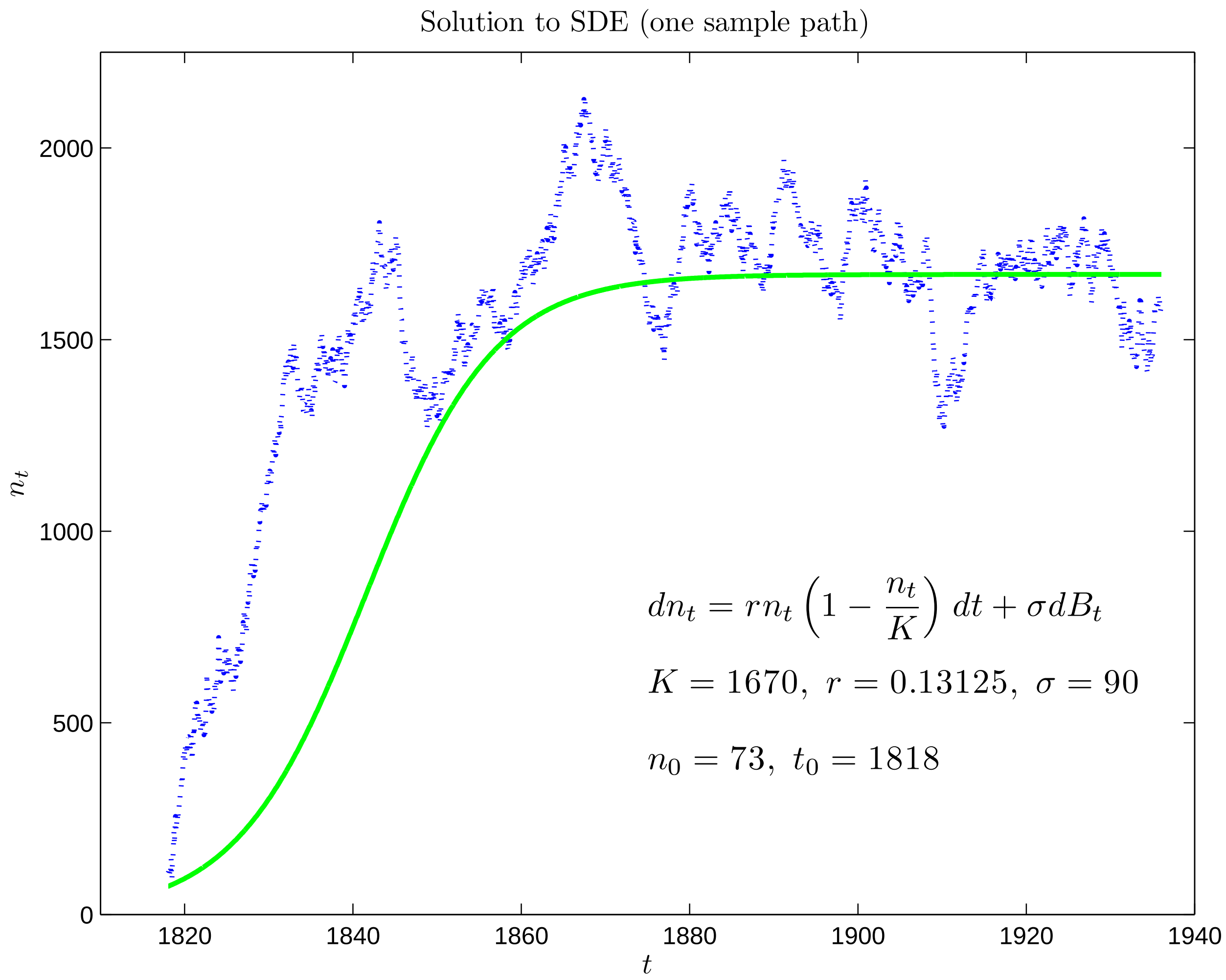


Solution to SDE (Run 1)



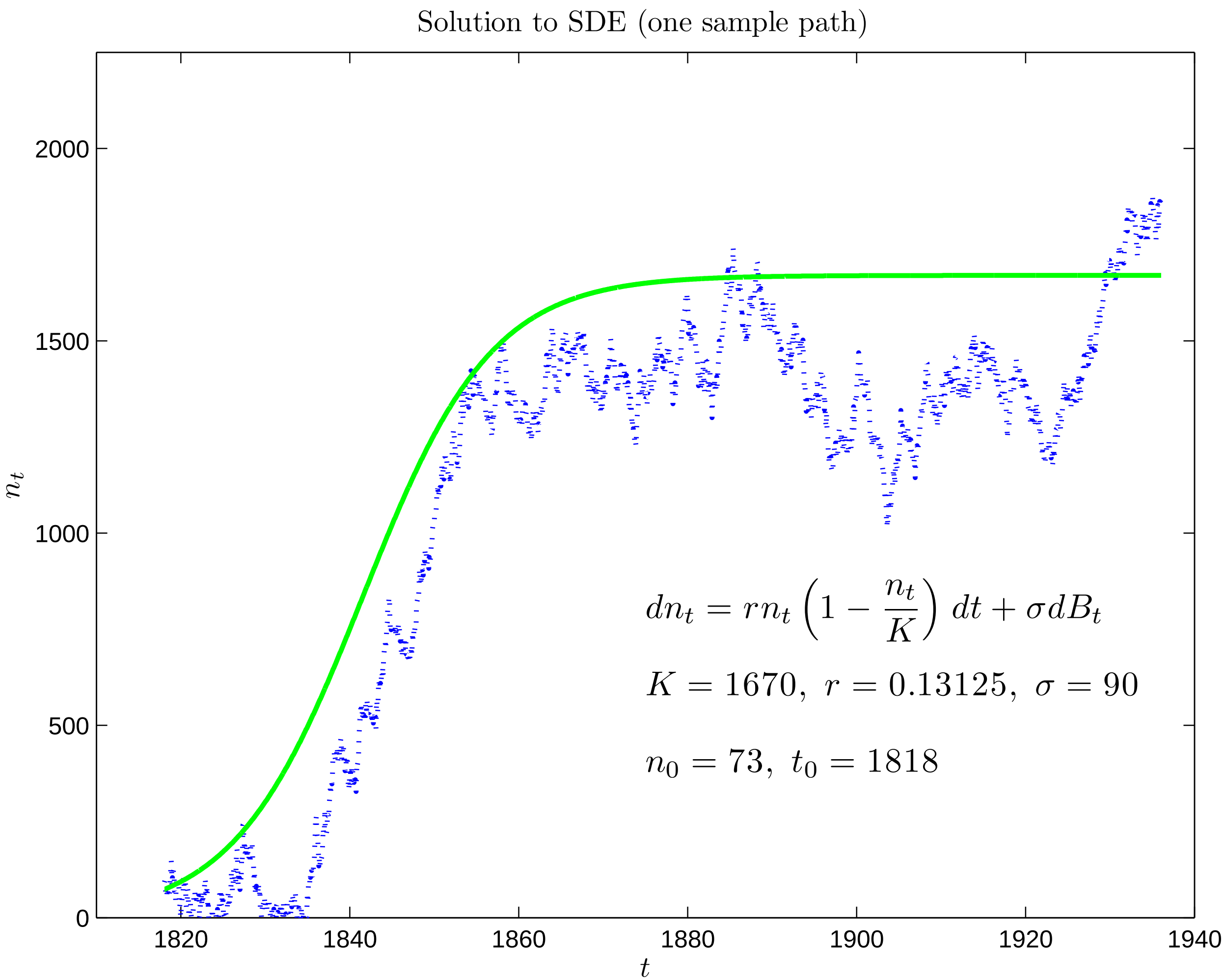
(Solution to the deterministic model is in green)

Solution to SDE (Run 2)



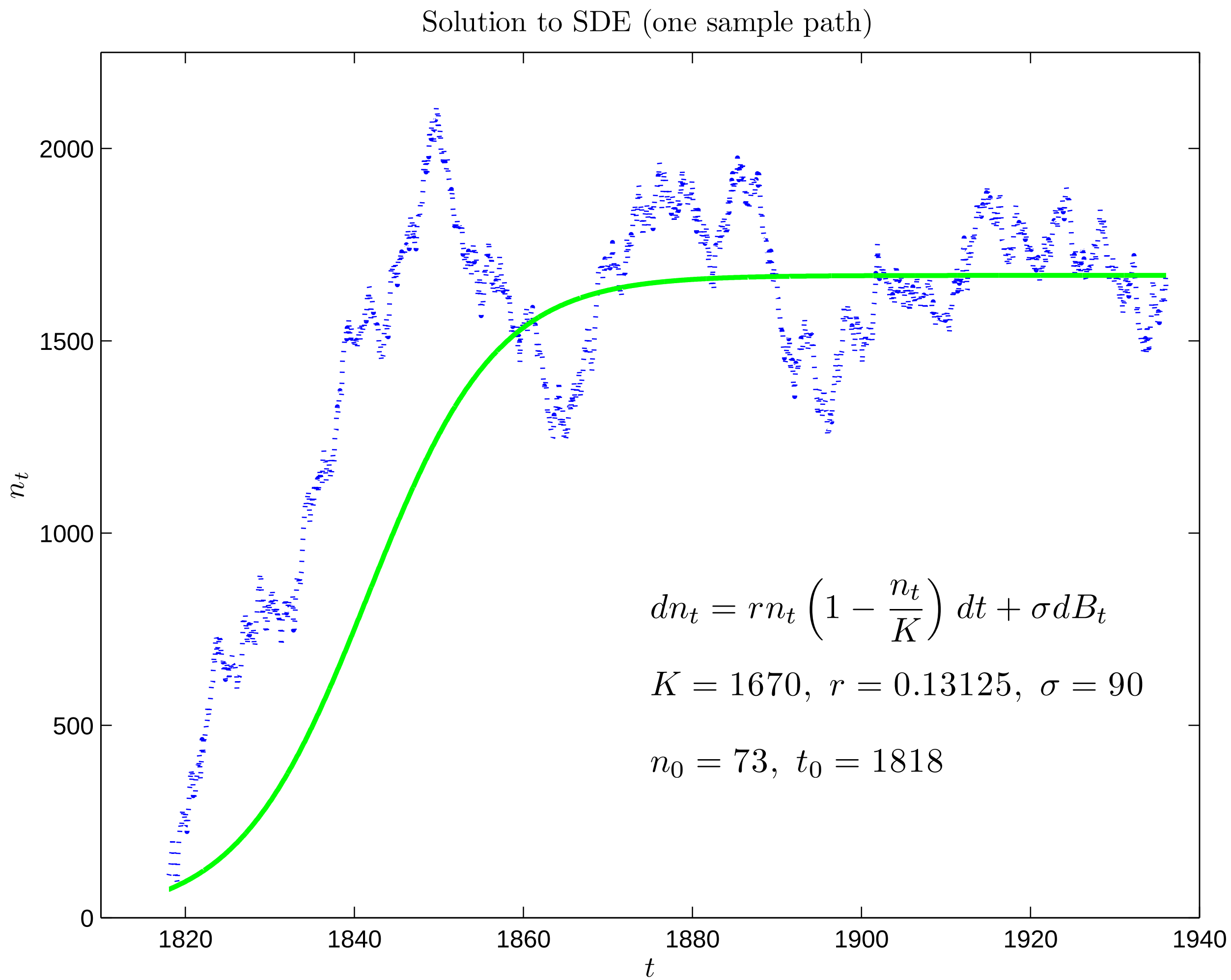
(Solution to the deterministic model is in green)

Solution to SDE (Run 3)



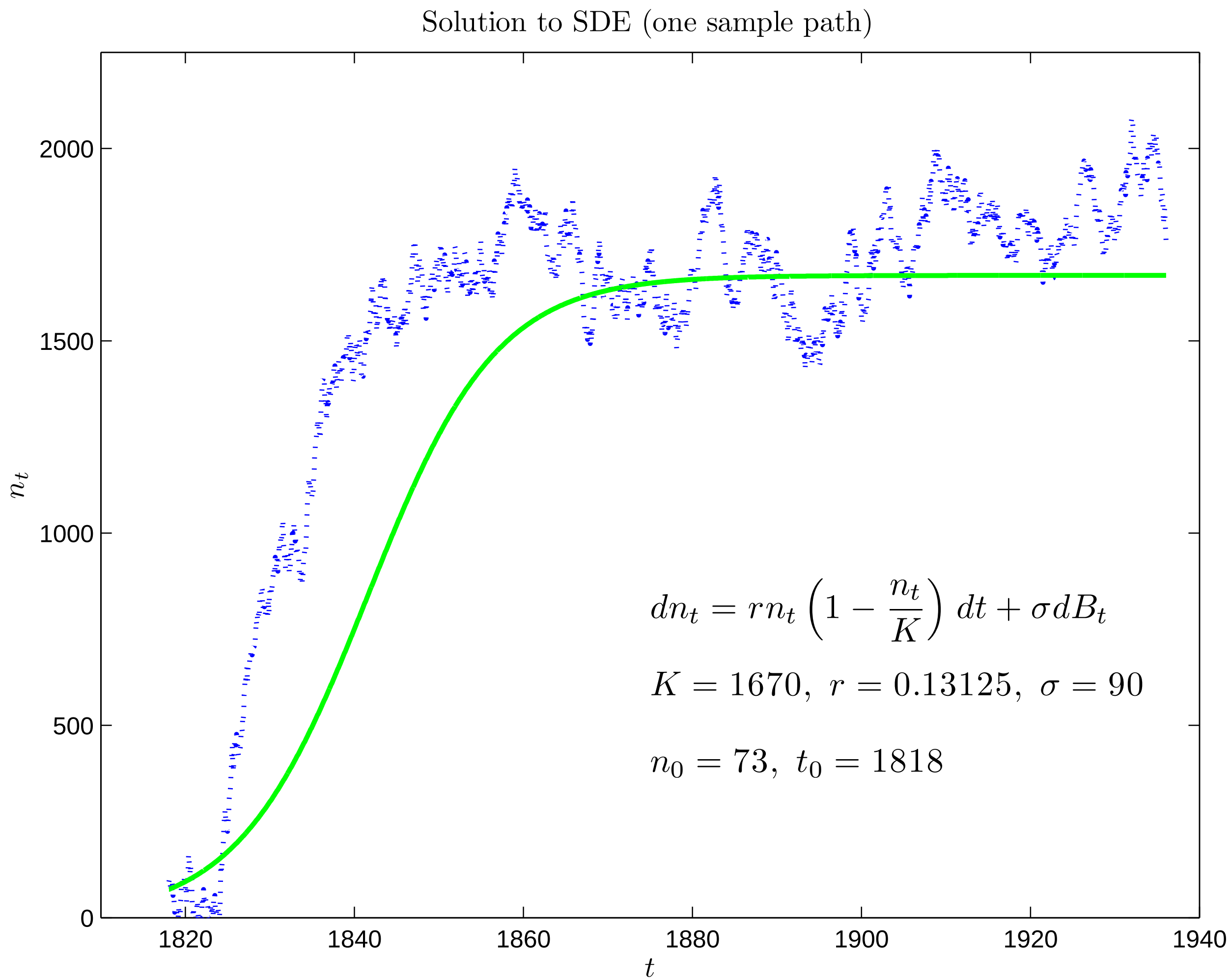
(Solution to the deterministic model is in green)

Solution to SDE (Run 4)



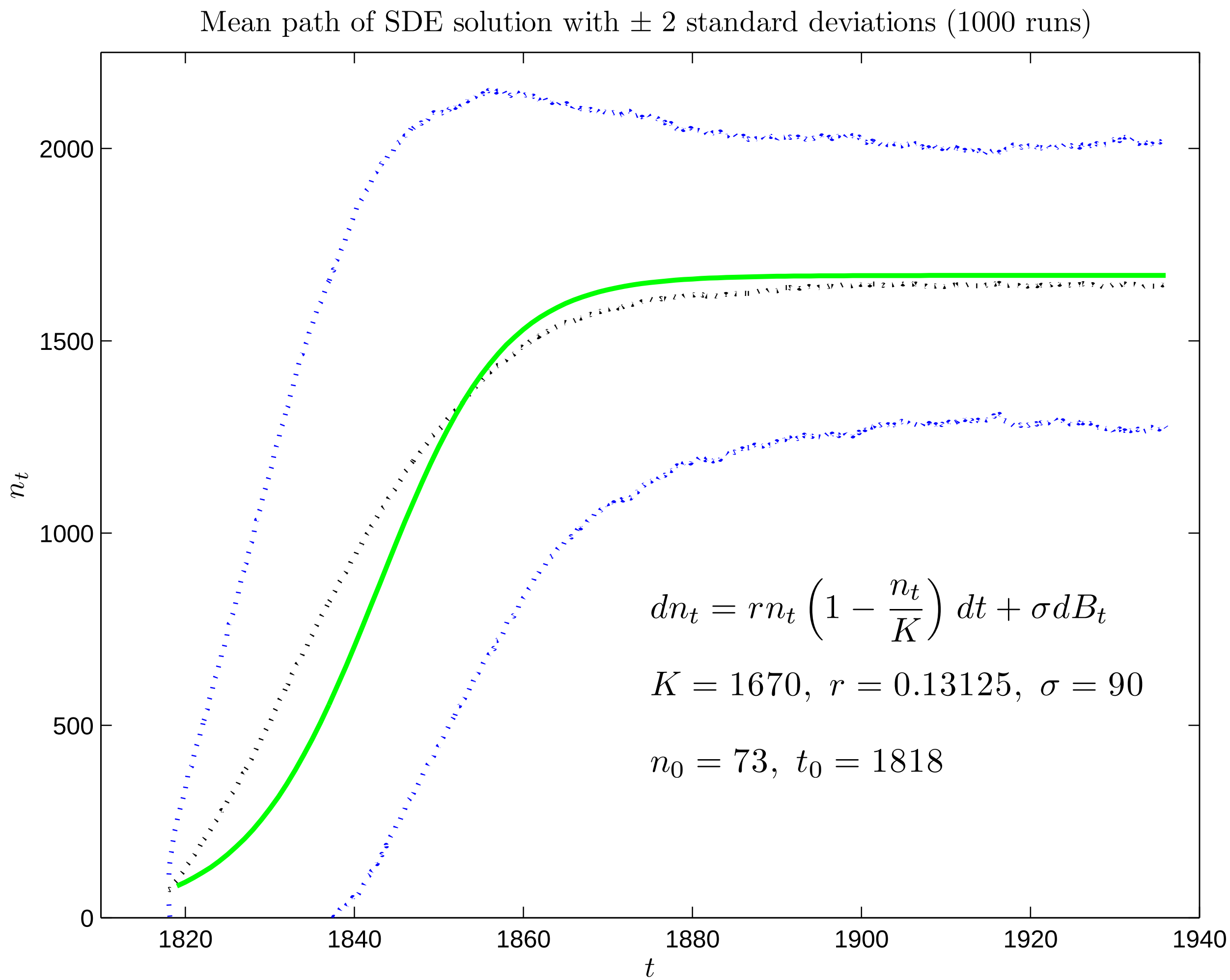
(Solution to the deterministic model is in green)

Solution to SDE (Run 5)



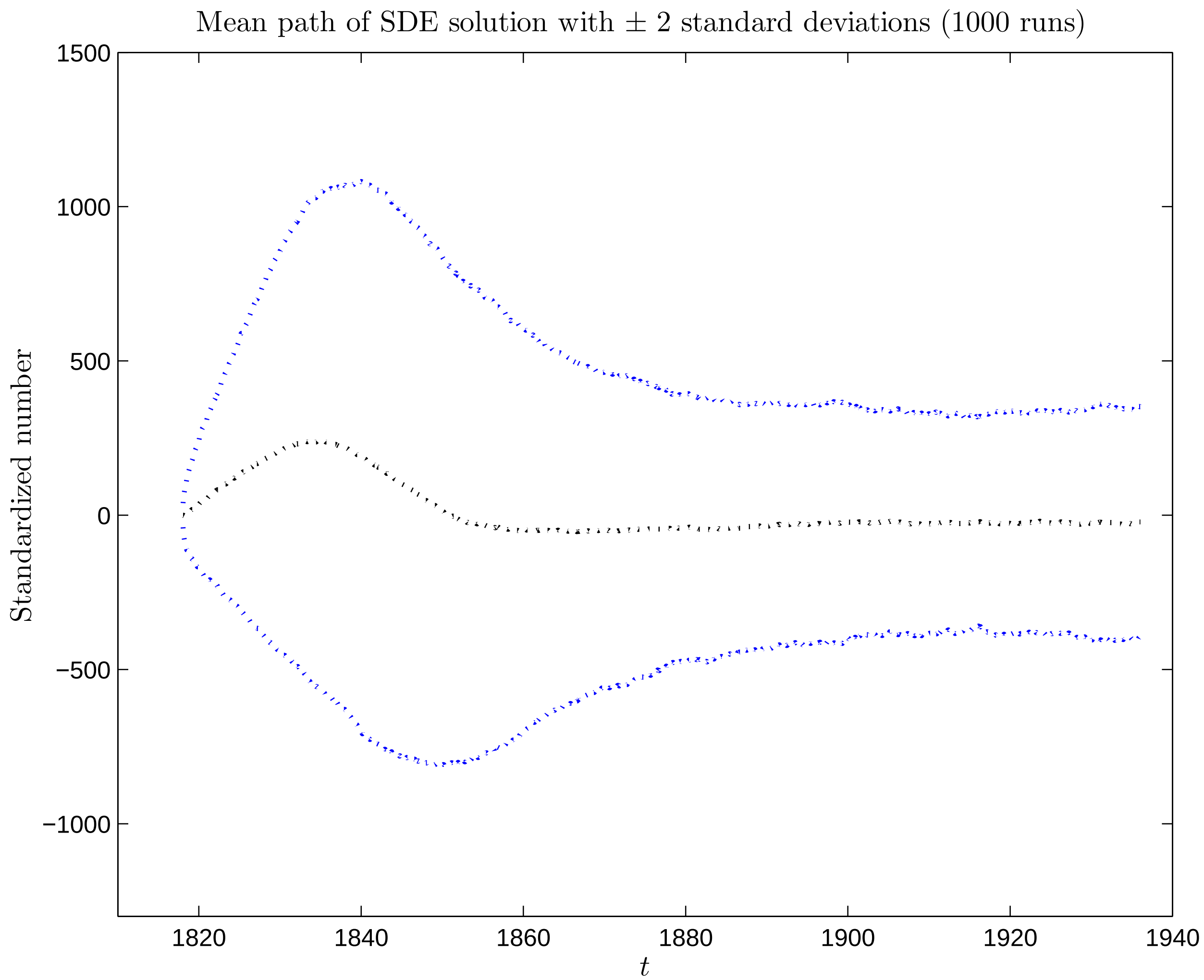
(Solution to the deterministic model is in green)

Solution to SDE



(Solution to the deterministic model is in green)

Solution to SDE



(With the solution to the deterministic model subtracted)

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t\left(1 - \frac{n_t}{K}\right)dt + \sigma\, dB_t.$$

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t\left(1 - \frac{n_t}{K}\right)dt + \sigma\, dB_t.$$

For a start:

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t \left(1 - \frac{n_t}{K}\right) dt + \sigma \, dB_t.$$

For a start:

- 0 is *reflecting*;

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t \left(1 - \frac{n_t}{K}\right) dt + \sigma \, dB_t.$$

For a start:

- 0 is *reflecting*;
- The mean path of the SDE solution does not follow a logistic curve;

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t\left(1 - \frac{n_t}{K}\right)dt + \sigma\, dB_t.$$

For a start:

- 0 is *reflecting*;
- The mean path of the SDE solution does not follow a logistic curve;
- The variance in the solution is large for the non-equilibrium phase–is this okay?

Logistic model with noise

So, what is wrong with the model?

$$dn_t = rn_t\left(1 - \frac{n_t}{K}\right)dt + \sigma\, dB_t.$$

For a start:

- 0 is *reflecting*;
- The mean path of the SDE solution does not follow a logistic curve;
- The variance in the solution is large for the non-equilibrium phase–is this okay?

$\ldots$ not to mention the fact that n_t is a continuous variable, yet population size is an integer-valued process!

The variance!

Since the variance is not uniform over time, we should *at least* have

$$dn_t = rn_t \left(1 - \frac{n_t}{K}\right) dt + \sigma(n_t)\, dB_t,$$

if not

$$dn_t = rn_t \left(1 - \frac{n_t}{K}\right) dt + \sigma(n_t, t)\, dB_t.$$

A different approach

Let's start from scratch specifying a stochastic model with variation being an inherent property: a *Markovian model*.

A different approach

Let's start from scratch specifying a stochastic model with variation being an inherent property: a *Markovian model*.

We will suppose that n_t (integer-valued!) evolves as a birth-death process with rates

$$q_{n,n+1} = \lambda n \left(1 - \frac{n}{N}\right) \quad \text{and} \quad q_{n,n-1} = \mu n,$$

where λ and μ (both positive) are per-capita birth and death rates (for λ when the population is small). Here N is the *population ceiling* (n_t now takes values in $S = \{0, 1, \ldots, N\}$).

I will call this model the *stochastic logistic (SL) model*, though it has many names, having been rediscovered several times since Feller proposed it in 1939.

A different approach

Let's start from scratch specifying a stochastic model with variation being an inherent property: a *Markovian model*.

We will suppose that n_t (integer-valued!) evolves as a birth-death process with rates

$$q_{n,n+1} = \lambda n \left(1 - \frac{n}{N}\right) \quad \text{and} \quad q_{n,n-1} = \mu n,$$

where λ and μ (both positive) are per-capita birth and death rates (for λ when the population is small). Here N is the *population ceiling* (n_t now takes values in $S = \{0, 1, \ldots, N\}$).

I will call this model the *stochastic logistic (SL) model*, though it has many names, having been rediscovered several times since Feller proposed it in 1939.

It shares an important property with the deterministic logistic model: that of *density dependence*.

Density dependence

The Verhulst-Pearl model $\frac{dn}{dt} = rn\left(1 - \frac{n}{K}\right)$ can be written

$$\frac{1}{N}\frac{dn}{dt} = r\frac{n}{N}\left(1 - \frac{N}{K}\frac{n}{N}\right).$$

The rate of change of n_t depends on n_t only through $\frac{n_t}{N}$.

So, letting $x_t = n_t/N$ be the "population density", we get

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{E}\right), \quad \text{where} \quad E = K/N.$$

This is a convenient space scaling. We could have set $x_t = n_t/A$, where A is habitat area, and then

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{DE}\right), \quad \text{where} \quad D = N/A.$$

Markovian models

Let $(n_t,\, t \geq 0)$ be a continuous-time Markov chain taking values in $S \subseteq \mathrm{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$. We identify a quantity N, usually related to the size of the system being modelled.

Definition (Kurtz*) The model is *density dependent* if there is a subset E of R^k and a continuous function $f : \mathrm{Z}^k \times E \to \mathrm{R}$, such that

$$q_{n,n+l} = N f_l\left(\frac{n}{N}\right), \quad l \neq 0 \quad (l \in \mathrm{Z}^k).$$

(So, the idea is the same: the rate of change of n_t depends on n_t only through the "density" n_t/N.)

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

Markovian models

Let $(n_t,\, t \geq 0)$ be a continuous-time Markov chain taking values in $S \subseteq \mathbb{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$. We identify a quantity N, usually related to the size of the system being modelled.

Definition (Kurtz*) The model is *density dependent* if there is a subset E of $\mathbb{R}^k$ and a continuous function $f : \mathbb{Z}^k \times E \to \mathbb{R}$, such that

$$q_{n,n+l} = N f_l \left(\frac{n}{N} \right), \quad l \neq 0 \quad (l \in \mathbb{Z}^k).$$

(So, the idea is the same: the rate of change of n_t depends on n_t only through the "density" n_t/N.)

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

Tom Kurtz

Thomas Kurtz (taken in 2003)

Density dependence

Consider the *forward equations* for $p_n(t) := \Pr(n_t = n)$. Let $q_n = \sum_{m \neq n} q_{nm}$. Then,

$$p_n'(t) = -q_n p_n(t) + \sum_{m \neq n} p_m(t) q_{mn},$$

Density dependence

Consider the *forward equations* for $p_n(t) := \Pr(n_t = n)$. Let $q_n = \sum_{m \neq n} q_{nm}$. Then,

$$p_n'(t) = -q_n p_n(t) + \sum_{m \neq n} p_m(t) q_{mn},$$

and so (formally) $\mathbb{E}(n_t) = \sum_n n p_n(t)$ satisfies

$$\frac{d}{dt}\mathbb{E}(n_t) = -\sum_n q_n n p_n(t) + \sum_m p_m(t) \sum_{n \neq m} n q_{mn}.$$

Density dependence

Consider the *forward equations* for $p_n(t) := \Pr(n_t = n)$. Let $q_n = \sum_{m \neq n} q_{nm}$. Then,

$$p_n'(t) = -q_n p_n(t) + \sum_{m \neq n} p_m(t) q_{mn},$$

and so (formally) $\mathbb{E}(n_t) = \sum_n n p_n(t)$ satisfies

$$\frac{d}{dt}\mathbb{E}(n_t) = -\sum_n q_n n p_n(t) + \sum_m p_m(t) \sum_{n \neq m} n q_{mn}.$$

So if $q_{n,n+l} = N f_l(n/N)$, then

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(n_t) &= -\sum_n \sum_{l \neq 0} N f_l(n/N) n p_n(t) \\ &\qquad + \sum_m p_m(t) \sum_{l \neq 0} (m + l) N f_l(m/N) \\ &= \sum_m p_m(t) N \sum_{l \neq 0} l f_l(m/N) = N\mathbb{E}\left(\sum_{l \neq 0} l f_l(n_t/N)\right).\end{aligned}$$

Density dependence

For an arbitrary density dependent model, define $F : E \rightarrow \mathrm{R}$ by $F(x) = \sum_{l \neq 0} l f_l(x)$. Then,

$$\frac{d}{dt}\,\mathbb{E}(n_t) = N\,\mathbb{E}\left(F\left(\frac{n_t}{N}\right)\right),$$

or, setting $X_t = n_t/N$ (the *density process*),

$$\frac{d}{dt}\,\mathbb{E}(X_t) = \mathbb{E}\left(F(X_t)\right).$$

Density dependence

For an arbitrary density dependent model, define $F : E \to \mathrm{R}$ by $F(x) = \sum_{l \neq 0} l f_l(x)$. Then,

$$\frac{d}{dt}\,\mathbb{E}(n_t) = N\,\mathbb{E}\left(F\left(\frac{n_t}{N}\right)\right),$$

or, setting $X_t = n_t/N$ (the *density process*),

$$\frac{d}{dt}\,\mathbb{E}(X_t) = \mathbb{E}\left(F(X_t)\right).$$

Warning: I'm not saying that $\frac{d}{dt}\,\mathbb{E}(X_t) = F\left(\mathbb{E}(X_t)\right)$.

Density dependence

For an arbitrary density dependent model, define $F : E \to \mathbb{R}$ by $F(x) = \sum_{l \neq 0} l f_l(x)$. Then,

$$\frac{d}{dt}\,\mathbb{E}(n_t) = N\,\mathbb{E}\left(F\left(\frac{n_t}{N}\right)\right),$$

or, setting $X_t = n_t/N$ (the *density process*),

$$\frac{d}{dt}\,\mathbb{E}(X_t) = \mathbb{E}\left(F(X_t)\right).$$

Warning: I'm not saying that $\frac{d}{dt}\,\mathbb{E}(X_t) = F\left(\mathbb{E}(X_t)\right)$.

(But, I *am* hoping for something like that to be true!)

Density dependence

For the SL model we have $S = \{0, 1, \ldots, N\}$ and

$$q_{n,n+1} = \lambda n \left(1 - \frac{n}{N}\right) \quad \text{and} \quad q_{n,n-1} = \mu n.$$

Therefore, $f_{+1}(x) = \lambda x\,(1 - x)$ and $f_{-1}(x) = \mu x$, $x \in E := [0, 1]$, and so $F(x) = \lambda x\,(1 - \rho - x)$, $x \in E$, where $\rho = \mu/\lambda$.

Density dependence

For the SL model we have $S = \{0, 1, \ldots, N\}$ and

$$q_{n,n+1} = \lambda n \left(1 - \frac{n}{N}\right) \quad \text{and} \quad q_{n,n-1} = \mu n.$$

Therefore, $f_{+1}(x) = \lambda x\,(1-x)$ and $f_{-1}(x) = \mu x$, $x \in E := [0,1]$, and so $F(x) = \lambda x\,(1-\rho-x)$, $x \in E$, where $\rho = \mu/\lambda$.

Now compare $F(x)$ with the right-hand side of the Verhulst-Pearl model for the density process:

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{E}\right), \quad \text{where} \quad E = K/N. \tag{2}$$

If $K \sim \beta N$ for N large, so that $K/N \to \beta$, then we may identify β with $1-\rho$ and r with $\lambda\beta$, and discover that (2) can be rewritten as $dx/dt = F(x)$.

Recall that ...

Recall that $(n_t,\, t \geq 0)$ is a continuous-time Markov chain taking values in $S \subseteq \mathrm{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$, and we have identified a quantity N, usually related to the size of the system being modelled.

The model is assumed to be *density dependent*: there is a subset E of R^k and a continuous function $f : \mathrm{Z}^k \times E \to \mathrm{R}$, such that

$$q_{n,n+l} = N f_l\left(\tfrac{n}{N}\right), \quad l \neq 0 \quad (l \in \mathrm{Z}^k).$$

We set $F(x) = \sum_{l \neq 0} l f_l(x)$, $x \in E$.

We now formally define the *density process* $(X_t^{(N)})$ by $X_t^{(N)} = n_t/N$, $t \geq 0$. We hope that $(X_t^{(N)})$ becomes more deterministic as N gets large.

Recall that ...

Recall that $(n_t,\, t \geq 0)$ is a continuous-time Markov chain taking values in $S \subseteq \mathrm{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$, and we have identified a quantity N, usually related to the size of the system being modelled.

The model is assumed to be *density dependent*: there is a subset E of R^k and a continuous function $f : \mathrm{Z}^k \times E \to \mathrm{R}$, such that

$$q_{n,n+l} = N f_l\left(\tfrac{n}{N}\right), \quad l \neq 0 \quad (l \in \mathrm{Z}^k).$$

We set $F(x) = \sum_{l \neq 0} l f_l(x)$, $x \in E$.

We now formally define the *density process* $(X_t^{(N)})$ by $X_t^{(N)} = n_t/N$, $t \geq 0$. We hope that $(X_t^{(N)})$ becomes more deterministic as N gets large. *To simplify the statement of results, I'm going to assume that the state space S is finite*.

A law of large numbers

The following *functional law of large numbers* establishes convergence of the family $(X_t^{(N)})$ to the unique trajectory of an appropriate approximating deterministic model.

Theorem (Kurtz*) Suppose F is Lipschitz on E (that is, $|F(x) - F(y)| < M_E|x-y|$). If $\lim_{N\to\infty} X_0^{(N)} = x_0$, then the density process $(X_t^{(N)})$ converges uniformly in probability on $[0,t]$ to (x_t), the unique (deterministic) trajectory satisfying

$$\frac{d}{ds}x_s = F(x_s), \quad x_s \in E,\ s \in [0,t].$$

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

A law of large numbers

The following *functional law of large numbers* establishes convergence of the family $(X_t^{(N)})$ to the unique trajectory of an appropriate approximating deterministic model.

Theorem (Kurtz*) Suppose F is Lipschitz on E (that is, $|F(x) - F(y)| < M_E|x - y|$). If $\lim_{N\to\infty} X_0^{(N)} = x_0$, then the density process $(X_t^{(N)})$ converges uniformly in probability on $[0,t]$ to (x_t), the unique (deterministic) trajectory satisfying

$$\frac{d}{ds}x_s = F(x_s), \quad x_s \in E,\ s \in [0,t].$$

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

(If S is an infinite set, we have the additional conditions $\sup_{x\in E} \sum_{l\neq 0} |l| f_l(x) < \infty$ and $\lim_{d\to\infty} \sum_{|l|>d} |l| f_l(x) = 0$, $x \in E$.)

A law of large numbers

Convergence *uniformly in probability* on $[0, t]$ means that for every $\epsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(\sup_{s\le t}\left|X_t^{(N)} - x_t\right| > \epsilon\right) = 0.$$

A law of large numbers

Convergence *uniformly in probability* on $[0, t]$ means that for every $\epsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(\sup_{s\le t} \left|X_t^{(N)} - x_t\right| > \epsilon\right) = 0.$$

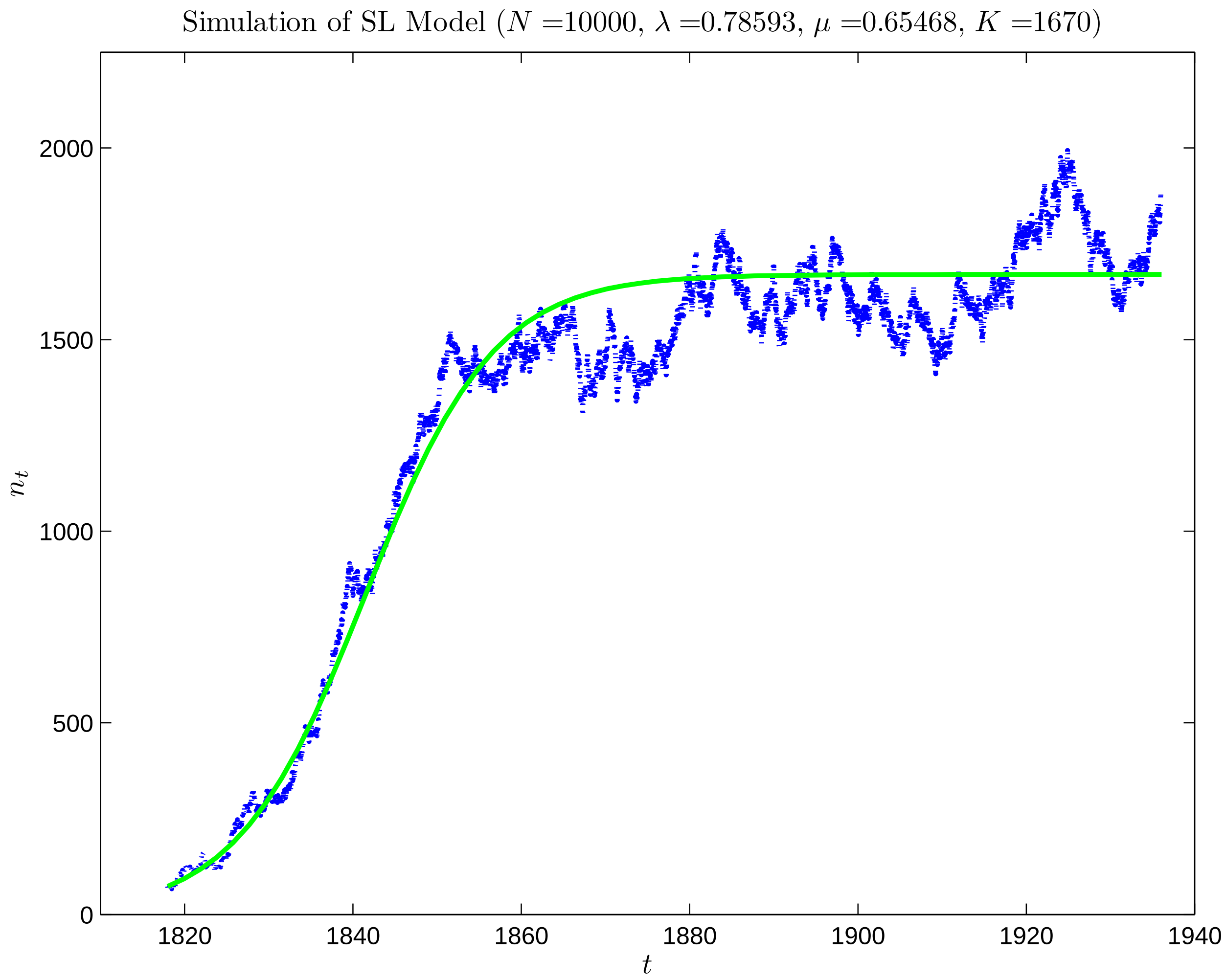
The conditions of the theorem hold for the SL model: since $F(x) = \lambda x(1 - \rho - x)$, we have, for all $x, y \in E = [0, 1]$, that

$$|F(x) - F(y)| = \lambda|x - y||1 - \rho - (x + y)| \le (1 + \rho)\lambda|x - y|.$$

So, provided $X_0^{(N)} \to x_0$ as $N \to \infty$, the population density $(X_t^{(N)})$ converges (uniformly in probability on finite time intervals) to the solution (x_t) of the deterministic model

$$\frac{dx}{dt} = \lambda x(1 - \rho - x) \qquad (x_t \in E).$$

Simulation of the SL model



(Solution to the deterministic model is in green)

A central limit law

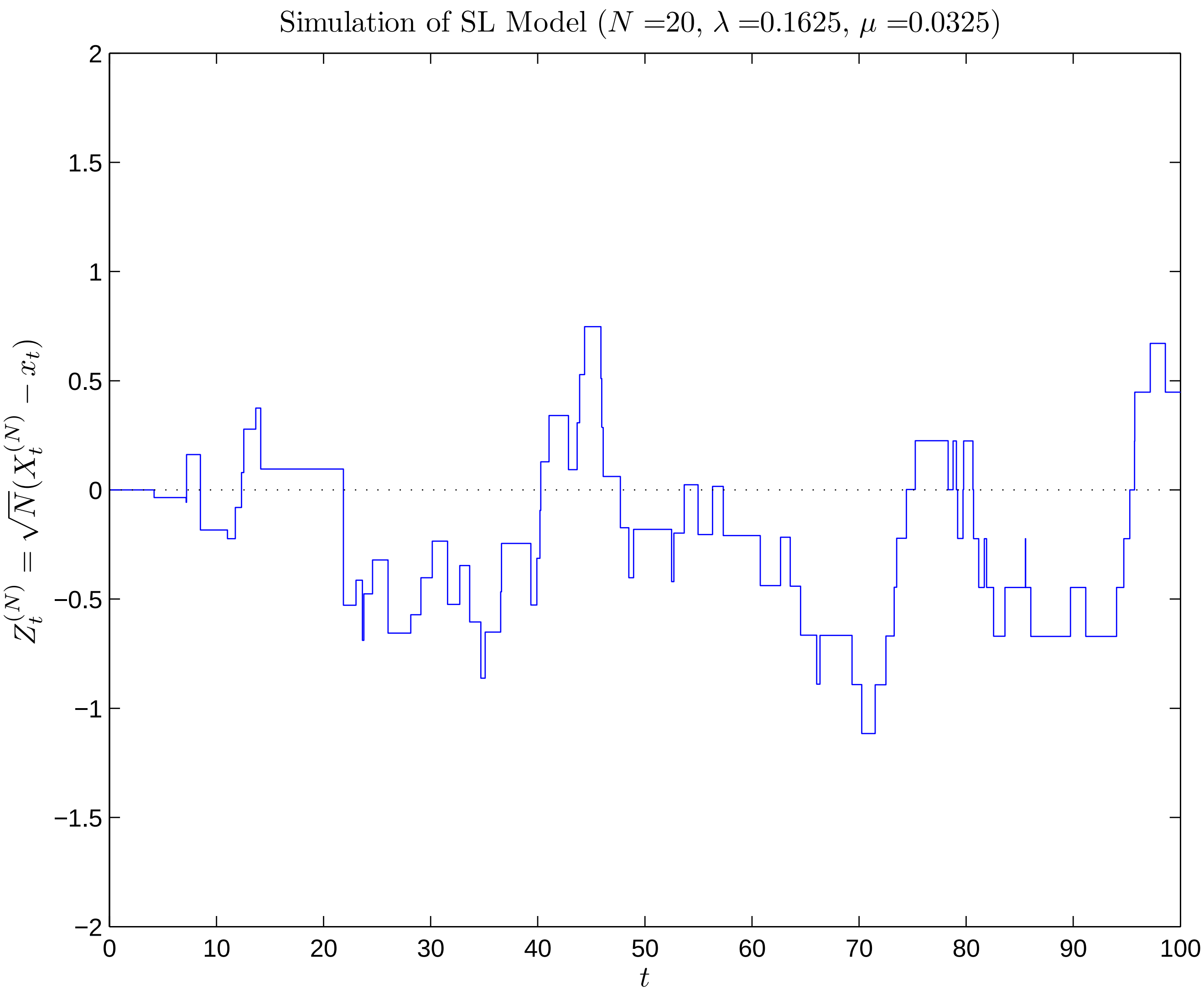
In a later paper Kurtz* proved a *functional central limit law* which establishes that, for large N, the fluctuations about the deterministic trajectory follow a *Gaussian diffusion*, provided that some mild extra conditions are satisfied.

He considered the family of processes $\{(Z_t^{(N)})\}$ defined by

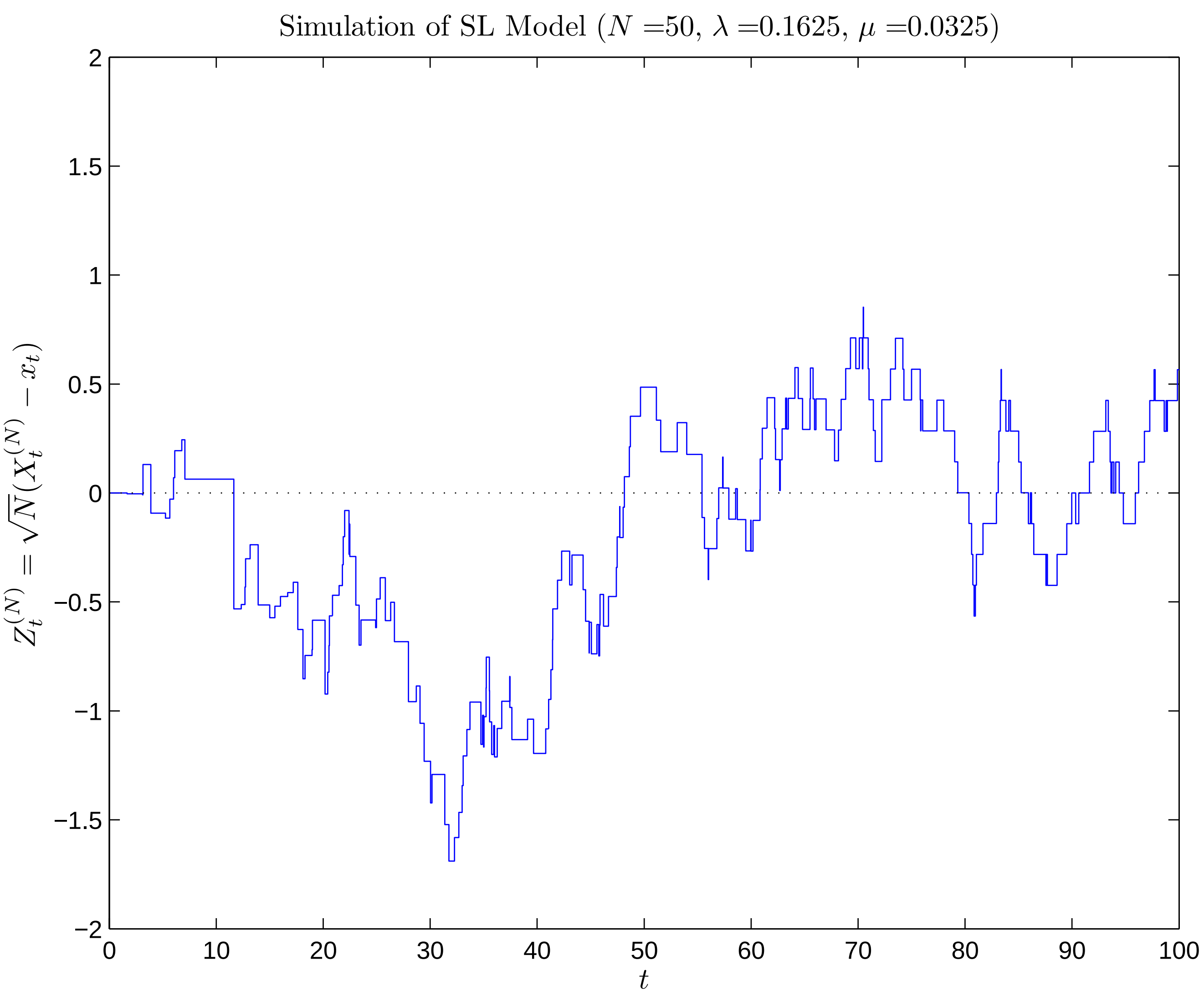
$$Z_s^{(N)} = \sqrt{N}\left(X_s^{(N)} - x_s\right), \qquad 0 \le s \le t.$$

*Kurtz, T. (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* 8, 344–356.

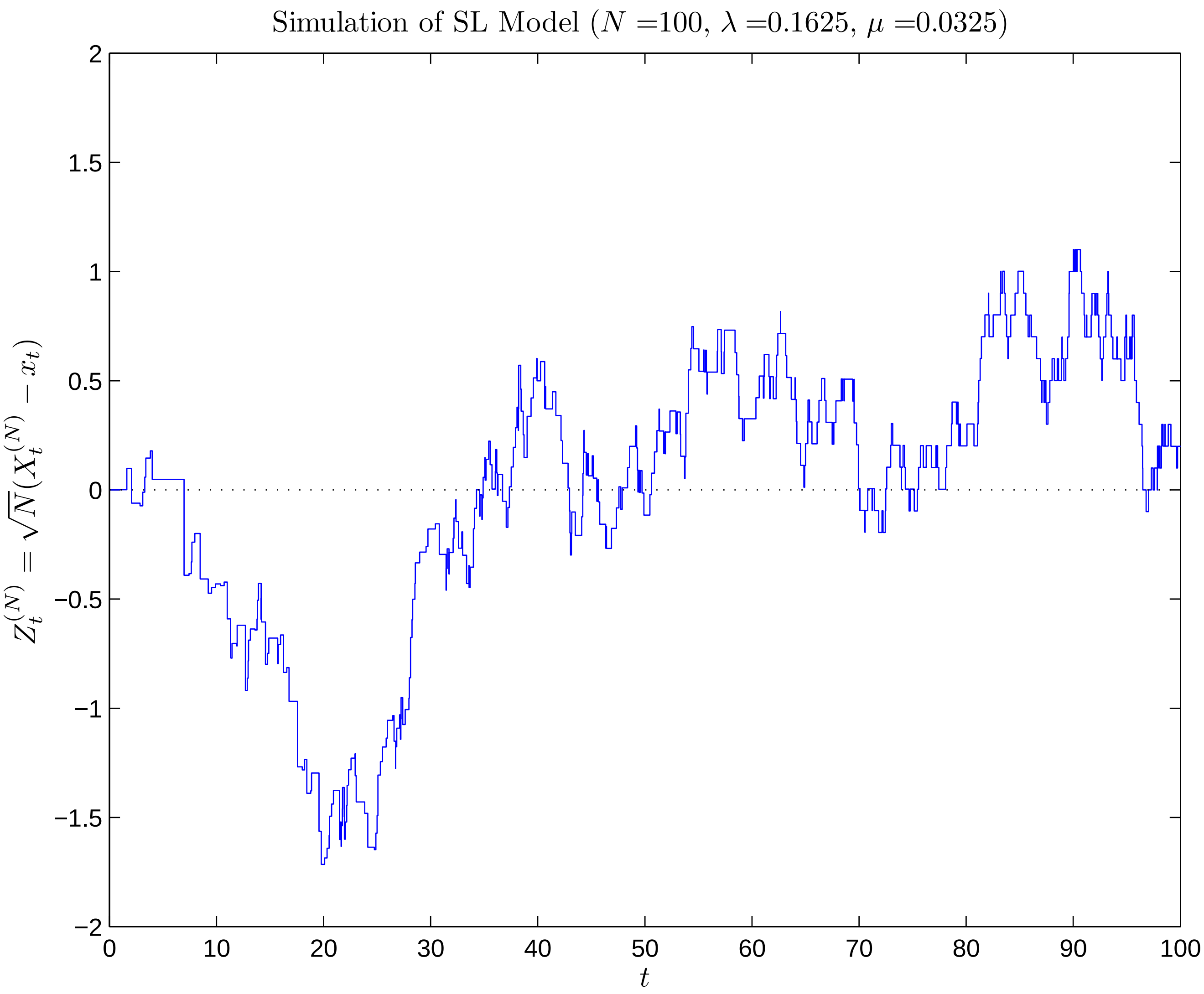
The SL model $(N = 20)$



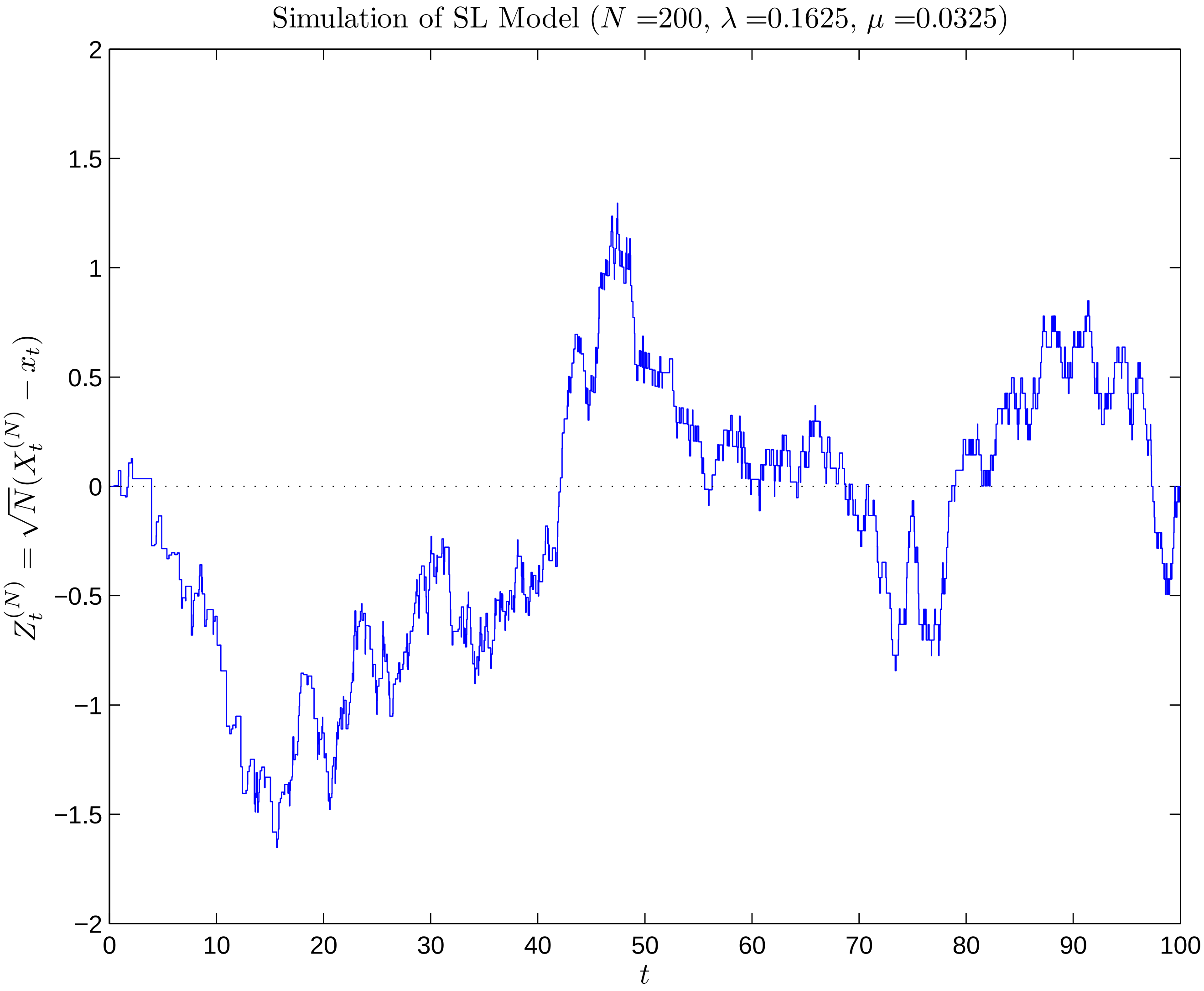
The SL model $(N = 50)$



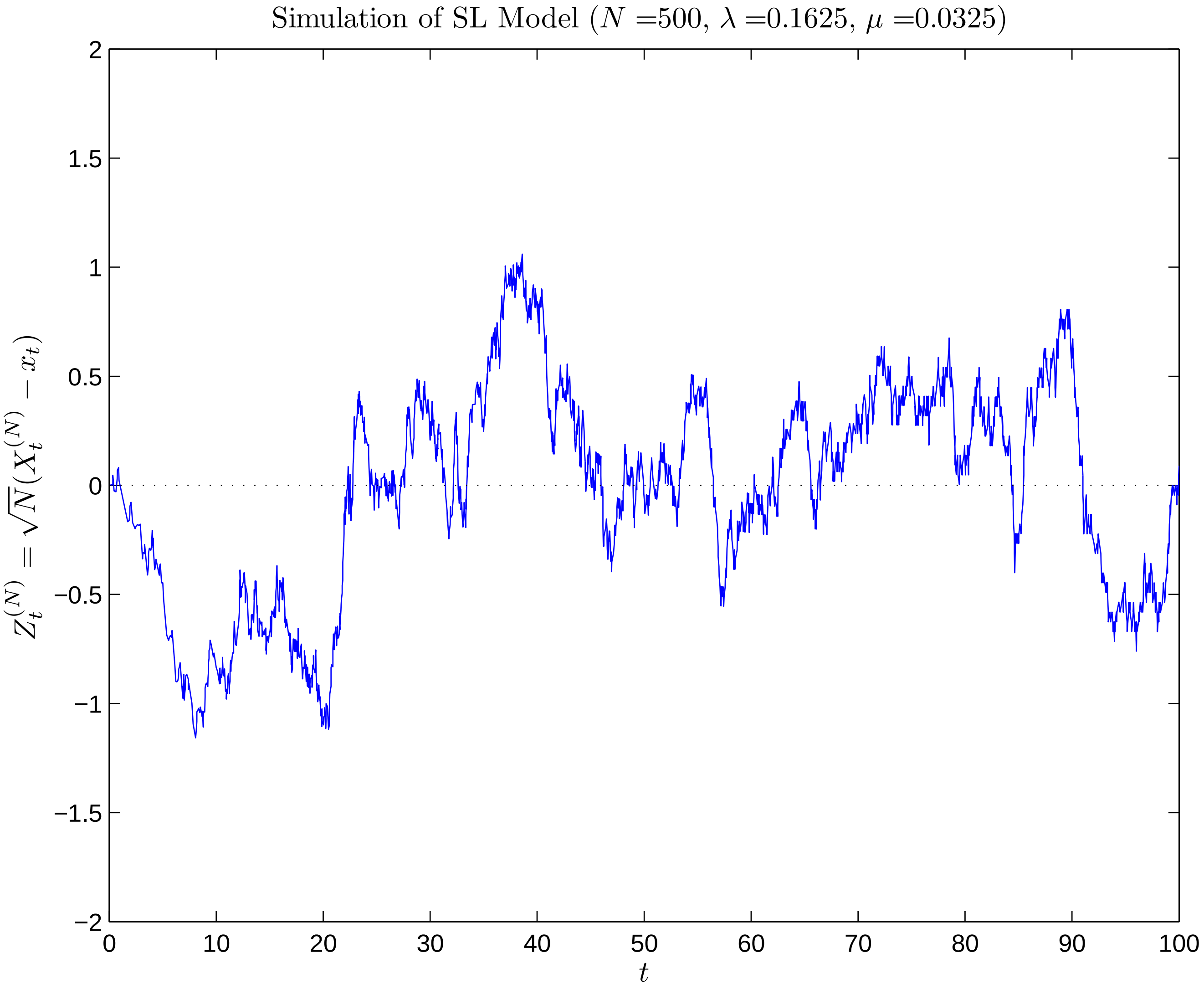
The SL model ($N = 100$)



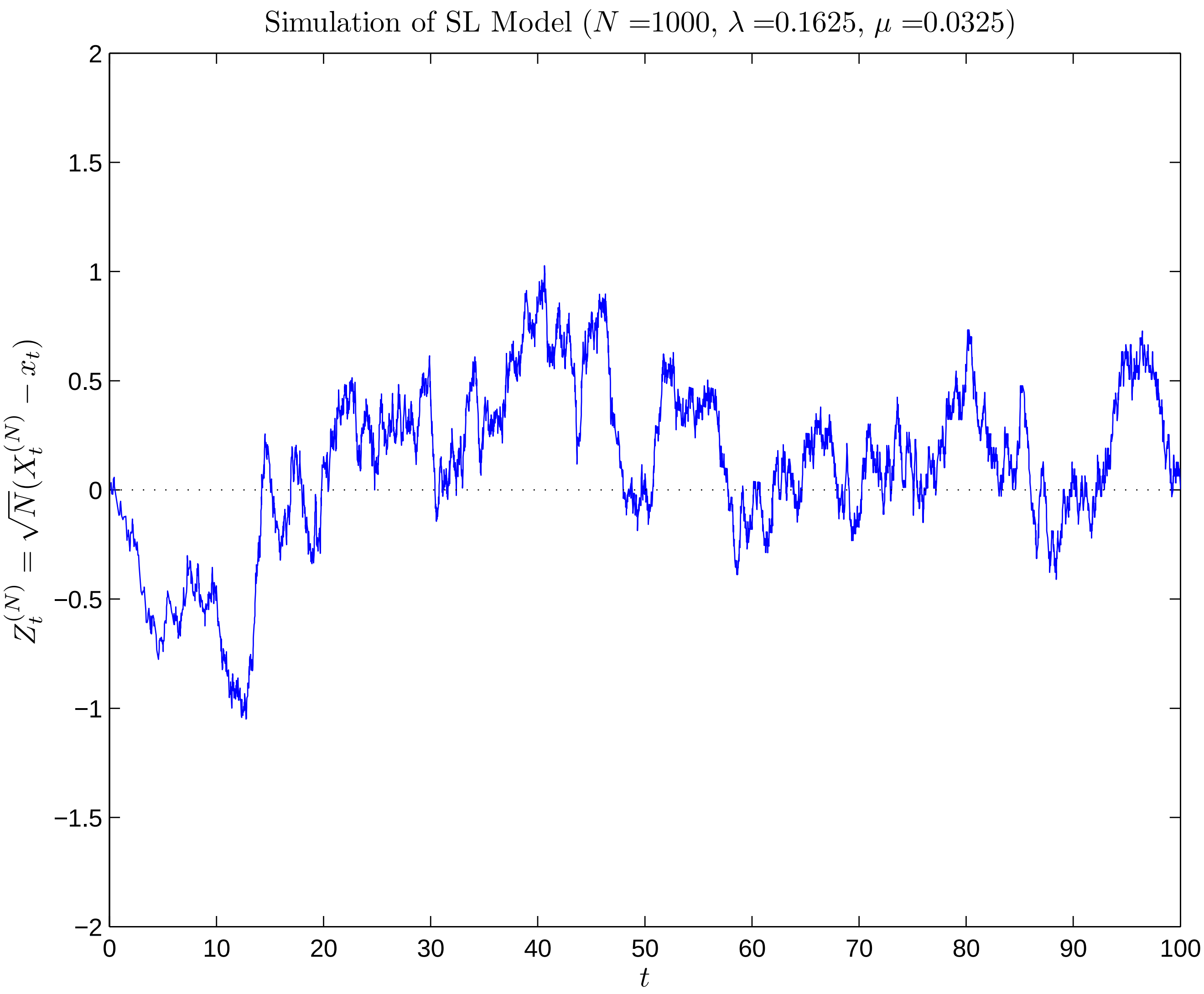
The SL model $(N = 200)$



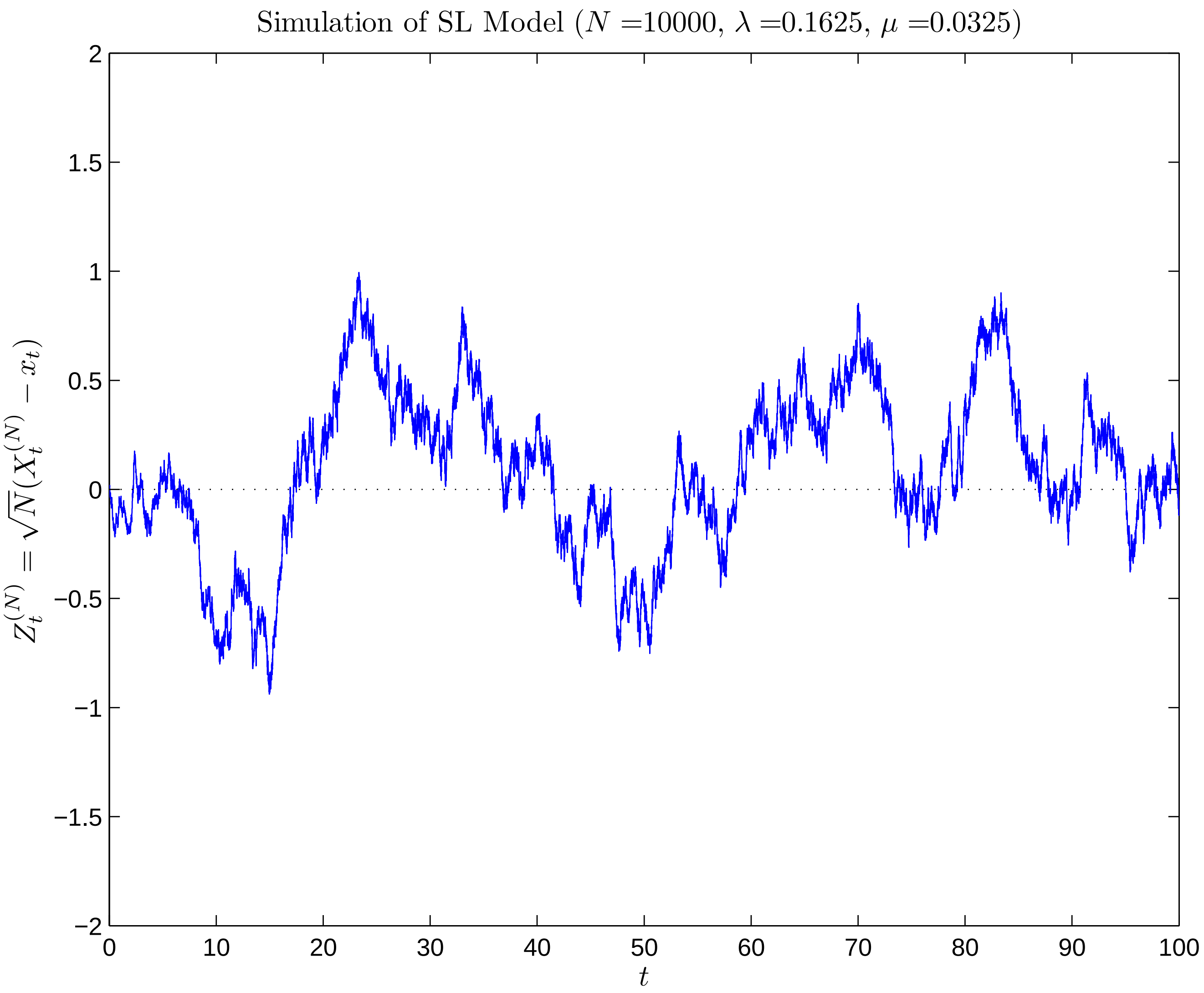
The SL model $(N = 500)$



The SL model $(N = 1\,000)$



The SL model $(N = 10\,000)$



A central limit law

Theorem Suppose that F is Lipschitz and has uniformly continuous first derivative on E, and that the $k \times k$ matrix $G(x)$, defined for $x \in E$ by $G_{ij}(x) = \sum_{l \neq 0} l_i l_j f_l(x)$, is uniformly continuous on E.

Let (x_t) be the unique deterministic trajectory starting at x_0 and suppose that $\lim_{N\to\infty} \sqrt{N}\left(X_0^{(N)} - x_0\right) = z$.

Then, $\{(Z_t^{(N)})\}$ converges weakly in $D[0,t]$ (the space of right-continuous, left-hand limits functions on $[0,t]$) to a Gaussian diffusion (Z_t) with initial value $Z_0 = z$ and with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = M_s z$, where $M_s = \exp(\int_0^s B_u \, du)$ and $B_s = \nabla F(x_s)$, and

$$V_s := \mathrm{Cov}(Z_s) = M_s \left(\int_0^s M_u^{-1} G(x_u) (M_u^{-1})^T \, du \right) M_s^T \, .$$

A central limit law

The functional central limit theorem tells us that, for large N, the scaled density process $Z_t^{(N)}$ can be approximated *over finite time intervals* by the Gaussian diffusion (Z_t).

In particular, for all $t > 0$, $X_t^{(N)}$ has an approximate normal distribution with $\mathrm{Cov}(X_t^{(N)}) \simeq V_t/N$.

We would usually take $x_0 = X_0^{(N)}$, thus giving $\mathbb{E}(X_t^{(N)}) \simeq x_t$.

A central limit law

For the SL model we have $F(x) = \lambda x(1-\rho-x)$, and the solution to $dx/dt = F(x)$ is

$$x(t) = \frac{(1-\rho)x_0}{x_0+(1-\rho-x_0)e^{-\lambda(1-\rho)t}}.$$

We also have $F'(x) = \lambda(1-\rho-2x)$ and

$$G(x) = \sum_l l^2 f_l(x) = \lambda x(1+\rho-x) = F(x) + 2\mu x,$$

giving

$$M_t = \exp\left(\int_0^t F'(x_s)\,ds\right) = \frac{(1-\rho)^2 e^{-\lambda(1-\rho)t}}{(x_0+(1-\rho-x_0)e^{-\lambda(1-\rho)t})^2}.$$

We can evaluate

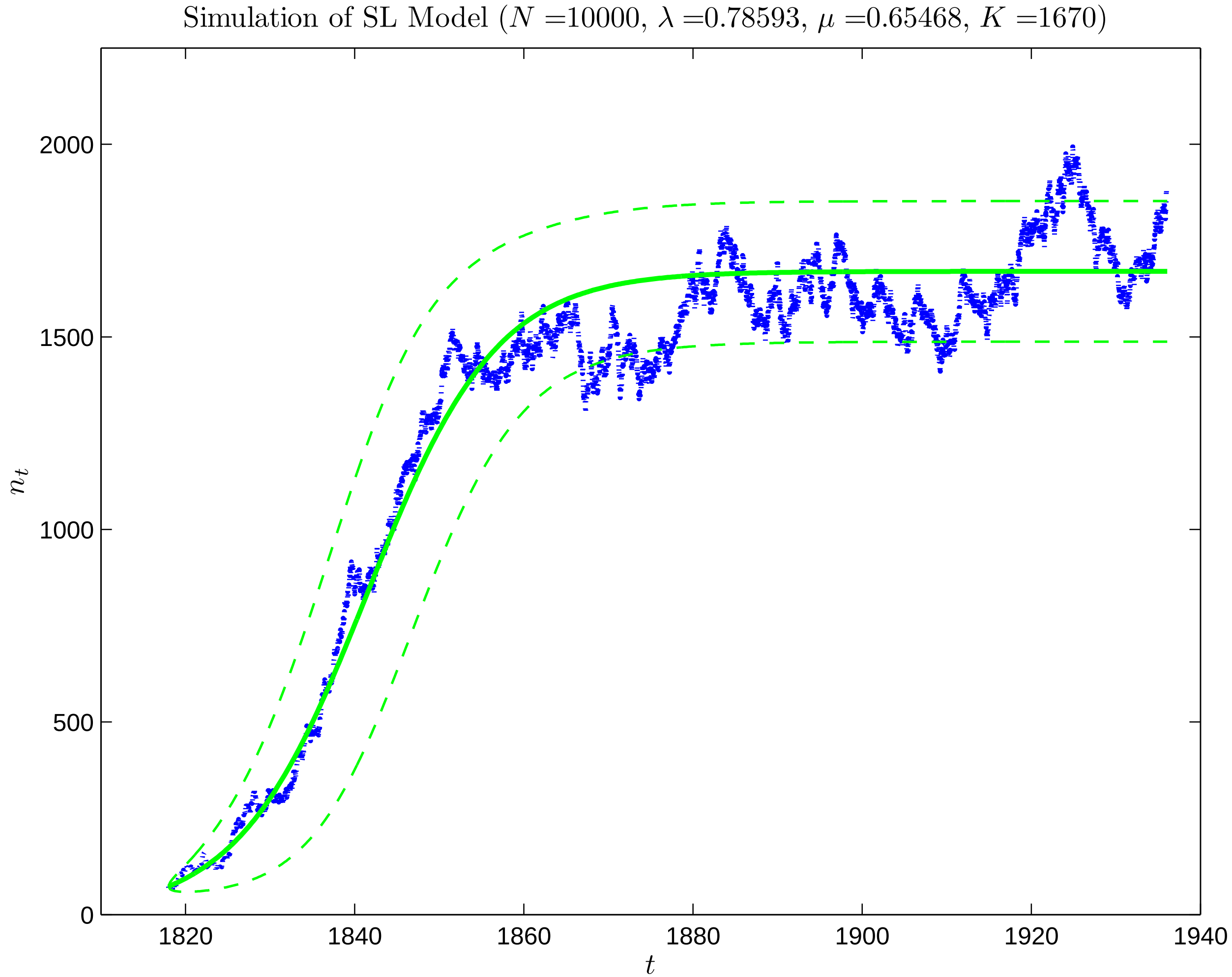
$$V_t := \mathrm{Var}(Z_t) = M_t^2\left(\int_0^t G(x_s)/M_s^2\,ds\right)$$

numerically, or ...

Or ....

$$
\begin{aligned}
V_t = x_0\Big(\rho x_0^3 + x_0^2(1+5\rho)(1-\rho-x_0)e^{-\lambda(1-\rho)t} \\
+ 2x_0(1+2\rho)(1-\rho-x_0)^2(\lambda(1-\rho)t)e^{-2\lambda(1-\rho)t} \\
- \big((1-\rho-x_0)[3\rho x_0^2 + (2+\rho)(1-\rho)x_0 - ((1+2\rho))(1-\rho)^2] \\
+ \rho(1-\rho)^3\big)e^{-2\lambda(1-\rho)t} \\
-(1+\rho)(1-\rho-x_0)^3 e^{-3\lambda(1-\rho)t}\Big)\Big/\Big(x_0+(1-\rho-x_0)e^{-\lambda(1-\rho)t}\Big)^4.
\end{aligned}
$$

The SL model



(Deterministic trajectory plus or minus two standard deviations in green)

The OU approximation

If the initial point x_0 of the deterministic trajectory is chosen to be an equilibrium point of the deterministic model, we can be far more precise about the approximating diffusion.

Corollary If $x_{\rm eq}$ satisfies $F(x_{\rm eq}) = 0$, then, under the conditions of the theorem, the family $\{(Z_t^{(N)})\}$, defined by

$$Z_s^{(N)} = \sqrt{N}(X_s^{(N)} - x_{\rm eq}), \qquad 0 \le s \le t,$$

converges weakly in $D[0,t]$ to an *OU process* (Z_t) with initial value $Z_0 = z$, local drift matrix $B = \nabla F(x_{\rm eq})$ and local covariance matrix $G(x_{\rm eq})$. In particular, Z_s is normally distributed with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = e^{Bs}z$ and

$$V_s := {\rm Cov}(Z_s) = \int_0^s e^{Bu} G(x_{\rm eq}) e^{B^T u}\, du\,.$$

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\rm eq}) e^{B^T u}\, du = V_\infty - e^{Bs} V_\infty e^{B^T s},$$

where V_∞, the stationary covariance matrix, satisfies

$$BV_\infty + V_\infty B^T + G(x_{\rm eq}) = 0.$$

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\rm eq}) e^{B^T u}\, du = V_\infty - e^{Bs} V_\infty e^{B^T s},$$

where V_∞, the stationary covariance matrix, satisfies

$$BV_\infty + V_\infty B^T + G(x_{\rm eq}) = 0.$$

We conclude that, for N large, $X_t^{(N)}$ has an approximate Gaussian distribution with $\mathrm{Cov}(X_t^{(N)}) \simeq V_t/N$.

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\rm eq}) e^{B^T u}\, du = V_\infty - e^{Bs} V_\infty e^{B^T s},$$

where V_∞, the stationary covariance matrix, satisfies

$$BV_\infty + V_\infty B^T + G(x_{\rm eq}) = 0.$$

We conclude that, for N large, $X_t^{(N)}$ has an approximate Gaussian distribution with $\mathrm{Cov}(X_t^{(N)}) \simeq V_t/N$.

For the SL model, we get $\mathrm{Var}(X_t^{(N)}) \simeq \rho(1 - e^{-2\lambda(1-\rho)t})/N$.

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\mathrm{eq}}) e^{B^T u}\, du = V_\infty - e^{Bs} V_\infty e^{B^T s},$$

where V_∞, the stationary covariance matrix, satisfies

$$BV_\infty + V_\infty B^T + G(x_{\mathrm{eq}}) = 0.$$

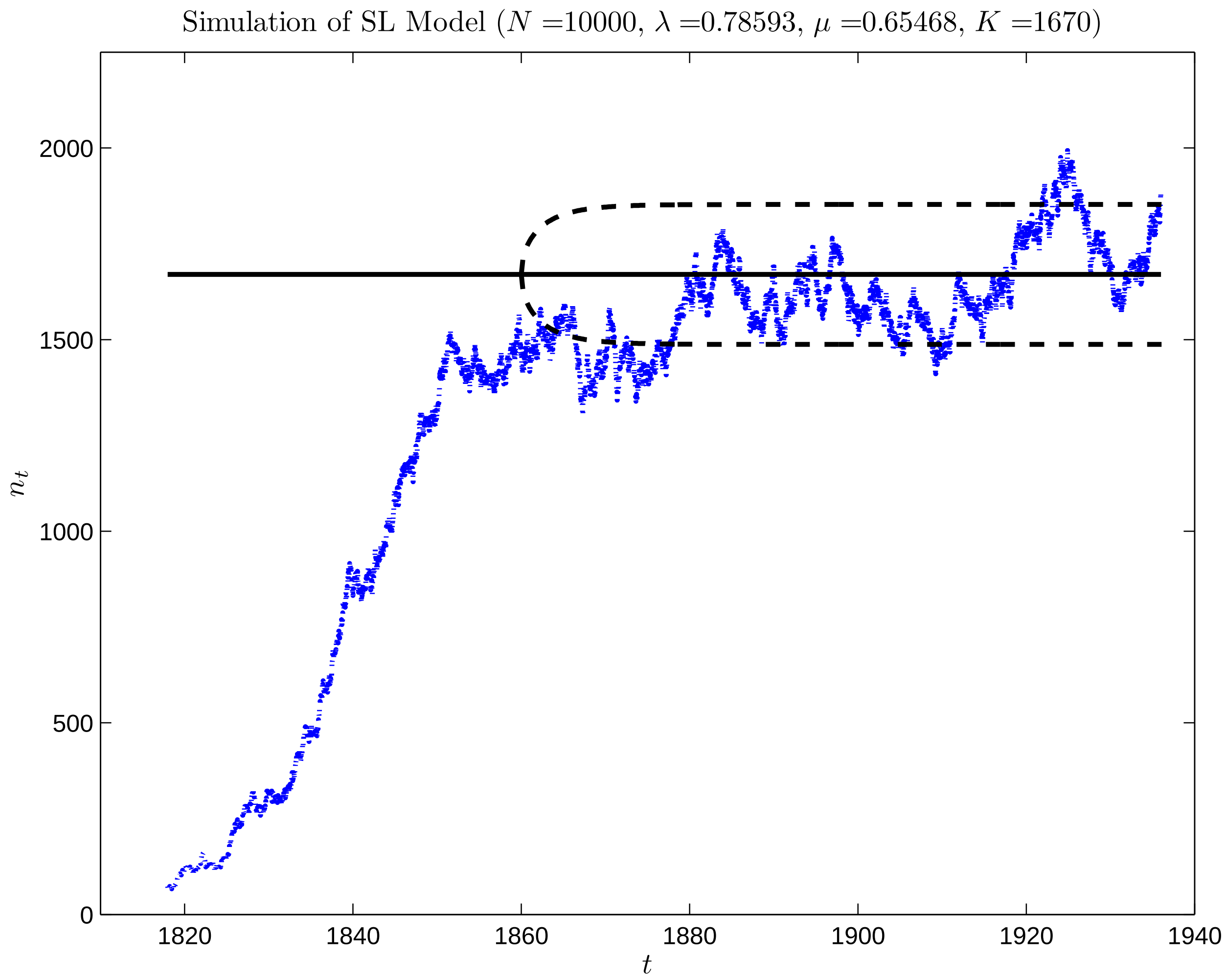
We conclude that, for N large, $X_t^{(N)}$ has an approximate Gaussian distribution with $\mathrm{Cov}(X_t^{(N)}) \simeq V_t/N$.

For the SL model, we get $\mathrm{Var}(X_t^{(N)}) \simeq \rho(1 - e^{-2\lambda(1-\rho)t})/N$.

This brings us “full circle” to the approximating SDE

$$dn_t = -\alpha(n_t - K)\, dt + \sqrt{2N\alpha\rho}\, dB_t, \quad \text{where } \alpha = \lambda(1-\rho).$$

The SL model



(Deterministic equilibrium plus or minus two standard deviations is in black)