Stochastic population modeling

Benito Chen-Charpentier

Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, USA
bmchen@uta.edu

African Mathematical School on Insight from Mathematical Modeling
into Problems in Conservation, Ecology, and Epidemiology
University of Cheikh Anta Diop, Dakar Senegal
May 2018

Outline

Introduction

An ordinary differential equation

$$\frac{dx}{dt} = f(x,t), \quad x(t_0) = x_0$$

can be written as

$$dx = f(x,t)dt, \quad x(t_0) = x_0$$

or in integral form

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s),s)\,ds.$$

A simple example is

$$\frac{dx}{dt} = a(t)x(t), \quad x(t_0) = x_0,$$

which is Malthus model for the growth of a population x.

Usually one considers that $a(t)$ is a deterministic parameter, but many times due to errors in measurement, variability in the populations and other factors that introduce uncertainties, we can think of $a(t)$ as a random variable, $a(t) = a_0(t) + h(t)\xi(t)$, where $a_0(t)$ is the expected value or mean, and $\xi(t)$ is a white noise process. (Later we will define these terms.) If we define $dW(t) = \xi(t)dt$ where $dW(t)$ is the differential form of the brownian motion, we get

$$dX(t) = a_0(t)X(t) + h(t)X(t)dW(t).$$

In general an stochastic differential equation is given by

$$dX(t,\omega) = f(X(t,\omega),t) + g(X(t,\omega),t)dW(t,\omega),$$

where ω is an element of the sample space and $X = X(t,\omega)$ is a random variable or stochastic process. Here we take the initial condition $X(0,\omega) = X_0$ to be known with probability one.

This equation is equivalent to the integral equation

$$X(t,\omega) = X_0 + \int_0^t f(X(s,\omega),s)\,ds + \int_0^t g(X(s,\omega),s)dW(s,\omega),$$

Later we will see what is the meaning of the stochastic integral, what is a brownian motion, etc.

Concepts of probability

- The result of a game or an experiment is many times random.
- For example, if we toss a coin, the possible results are "heads" (H) or "tails" (T). They are not predictable
- The sample space Ω is $\Omega = \{H, T\}$.
- The events are the subsets of Ω, $\{H\}, \{T\}, \{H, T\}$ and $\emptyset$.
- If the coin is not loaded, the probability of a heads is $1/2$ and the probability of a tails is also $1/2$.
- That is, if we toss the coin many times we expect to get heads half the time.

Another example: A day may be rainy (R), snowy (S) or clear (C)

- The sample space Ω is $\Omega = \{R, S, C\}$.
- The events are the subsets of Ω, $\{R\}, \{S\}, \{C\}, \{R, S\}, \{R, C\}, \{S, C\}, \{R, S, C\}$ and $\emptyset$.
- The probability of each happening depends on observations

Given an event $A \in \Omega$, the probability of A, $P(A)$, must satisfy:

1. $P(A) \in [0,1]$.
2. $P(\emptyset) = 0$.
3. $P(\Omega) = 1$.
4. IS $A_1, A_2, \ldots$ is a finite or countable sequence of disjoint events, then $P(A_1 \bigcup A_2 \ldots) = P(A_1) + P(A_2) + \cdots$

A random variable (r.v.) is a function that associates random outcomes ω in sample space Ω with scalars, usually a subset of the real numbers. E.g. associate a coin flip resulting in heads with 0, and tails with 1, or any other two numbers. Also, there are many other possible random variables associated with the same experiment.

- In an experiment or observation there is a probability space, which is an ordered triple $(\Omega, \mathscr{A}, P)$, with
  - Ω the sample space,
  - $\mathscr{A}$ a collection of subsets of of Ω and
  - P the probability measure defined on $\mathscr{A}$.

More precisely, the probability space is an ordered triple formed by:

- sample space Ω,
- σ-algebra $\mathscr{A}$ defined on Ω, and
- probability measure P defined on $\mathscr{A}$.

Given a nonempty set Ω, a σ-*algebra defined on* Ω, $\mathscr{A}$, is a collection of subsets of Ω such that

i) $\emptyset$ and Ω are both in $\mathscr{A}$,

ii) if A is in $\mathscr{A}$, then A^c is also in $\mathscr{A}$, where A^c denotes the complement of A, and

iii) if A_i is in $\mathscr{A}$ for each natural number i, then $\bigcup_{i\in\mathbb{N}} A_i$ is in $\mathscr{A}$.

A *non-negative probability measure* is a real map P from a σ-algebra $\mathscr{A}$ to [0,1] such that

i) $P(B) \geq 0$ for all $B \in \mathscr{A}$
ii) $P(\Omega) = 1$
iii) If $B_i \bigcap B_j = \emptyset$ for $1, j = 1, 2, \ldots, i \neq j$ then $P(\bigcup_{i \in \mathbb{N}} B_i) = \sum_{i \in \mathbb{N}} P(B_i)$, where each $B_i \in \mathscr{A}$

The *Borel σ-algebra* $\mathscr{G}$ is the smallest σ-algebra that contains all open sets in $\mathbb{R}^n$. Because of the way a σ-algebra is defined, $\mathscr{G}$ also contains all closed sets in $\mathbb{R}^n$, countable union of closed sets, countable intersection of open sets, and so on.
The measure μ_G defined on $\mathscr{G}$ given by

$$\mu_G(\prod_{i=1}^{n}[a_i, b_i]) = \prod_{i=1}^{n}(b_i - a_i)$$

is called the *Borel measure*

Restating: The ordered triple $(\Omega, \mathscr{B}, P)$ is called a *probability space*.

- Ω is the space in which the probability outcomes are found, called the sample space.
- $\mathscr{B}$ is a σ-algebra on Ω, and the elements of $\mathscr{B}$ are called *events*.
- P is called a *probability measure*, and it is a nonnegative measure with the following additional properties:
  i) $0 \leq P(A) \leq 1$ for $A \in \mathscr{B}$, and
  ii) $P(\Omega) = 1$.

- A *random variable* X is a map from Ω to $\mathbb{R}$, such that it is $\mathscr{B}$-measurable. That is, a random variable X is a map with the property that for each x in $\mathbb{R}$, the set $\{\omega \in \Omega : X(\omega) \leq x\}$ is in $\mathscr{B}$.
- Given an experiment with its associated random variable X and given a real number x, the probability of the event $\{\omega \in \Omega : X(\omega) \leq x\}$ is defined as the cumulative distribution function of X, and is denoted by

$$F_X(x) = P(\omega \in \Omega : X(\omega) \leq x) = P(X \leq x).$$

- If the random variable is continuous, then the derivative

$$f_X(x) = \frac{dF_X(x)}{dx}$$

is known as the probability density function of the random variable X, if the derivative exists.

If the random variable X has a probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

we say that the variable X has a normal distribution function with mean μ and variance σ^2, and it is denoted by

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

A stochastic process is a collection of parameterized random variables, usually the parameter is time. It is written as $\{X(t,\omega) : t \in T, \omega \in \Omega\}$. For t fixed, $X(t,\omega)$ is a random variable and fixing $\omega \in \Omega$, $X(t,\omega)$ it is a deterministic function of t that we will call a trajectory or a realization of the stochastic process $X(t,\omega)$.
Usually it is just written as $X(t)$.

To get information about the variability of a random variable or of a stochastic process one can use the moments, especially the first and second. The nth moment of a continuous variable X is

$$E(X^n) = \int_{\mathbb{R}} x^n f_X(x)\, dx, \text{if the integral exists.}$$

The mean is the first moment

$$E(X) = \int x dF(x),$$

and the variance is the second central moment

$$Var(X) = \int (x - E(X))^2 dF(x) = E((X - E(X)^2) = E(X^2) - E(X)^2.$$

The integrals are over the sample space.

The conditional expectation of a r.v. X given σ-algebra $\mathscr{C}$ denoted by $E(X|\mathscr{C})$ is the best approximation of X with respect to sets in $\mathscr{C}$.
In the special case X is $\mathscr{C}$-measurable,

$$E(X|\mathscr{C}) = X \text{ with probability 1},$$

and in case X is independent of sets in $\mathscr{C}$,

$$E(X|\mathscr{C}) = E(X) \text{ w.p.1.}$$

Assume that $\{X(t) : t \in [0,\infty)\}$ is a continuous in time stochastic process. Then we say that it is a Markov process if, given any sequence of times $0 < t_1 < t_2, \cdots < t_n$,

$$P(X(t_n) \leq y | X(0) = x_0, X(t_1) = x_1, \ldots, X(t_{n-1}) = x_{n-1}) =$$
$$P(X(t_n) \leq y | X(t_{n-1}) = x_{n-1}).$$

Uncertainty and variability in in physical, biological, social or economic phenomena can be modeled using stochastic processes. A class of frequently used stochastic processes is the Brownian Motion or Wiener process.

- First used to model the irregular movement of pollen on the surface of water, or of dust suspended in the air.
- Has been used to model different random phenomena where the next state depends only on the current state.
- Einstein explained it using the kinetic theory of gases.
- Wiener gave a mathematical formulation and therefore it is also known as a Wiener process.
- Mathematically it is formed by a sequence of random variables parameterized by time that are independent and identically distributed (iid).

Because of the independence assumption, the Brownian Motion is a convenient tool in stochastic modeling.

A stochastic process $W(t), t \in [0, \infty]$ is a Wiener process (or a standard Brownian motion) if it satisfies:

- It is defined for $t \geq 0$ with $W_0 = 0$.
- If $0 \leq s < t < \infty$, then $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$, that is $W(t) - W(s) \sim N(0, t - s)$.
- If $0 \leq r < s < t < \infty$, the increments $W(t) - W(s)$ and $W(s) - W(r)$ are independent.
- Since $\Delta W(t) = W(t + \Delta t) - W(t) \sim N(0, \Delta t)$ one has that $E(\Delta W(t))) = 0$ and $E((\Delta W(t))^2 = \Delta t$.

- Since $E((\Delta W(t))^2 = \Delta t$, $|\Delta W(t)|$ is $O(\Delta t^{1/2})$ the trajectories of a Wiener process are continuous: $\Delta t \to 0$ implies $(\Delta t)^{1/2} \to 0$.
- But
$$\frac{dW(T)}{dt} = \lim_{\Delta t \to 0} \frac{\Delta W(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{O(\Delta t^{1/2})}{\Delta t}$$
that does not exist.
- The Wiener process also does not have bounded variation and therefore one can not use integrals like Riemann-Stieltjes.

Ito Calculus

There are many ways to define stochastic integrals. The most common ones are Itô and Stratonovich integrals. Both make sense when integrating with respect to Wiener processes, but Itô are more widely used in biological applications.
Itô integral is defined in a similar way to Riemann or Riemann-Sieltjes integrals. Let $f(t)$ be a function of a random variable that satisfies $\int_a^b E(f^2(t)\,dt < \infty$. Let $a = t_1 < t_2 < \ldots < t_{n+1} = b$ be a partition of $[a, b]$ with $\Delta t = t_{i+1} - t_i = (b-a)/n$, and $\Delta W(t_i) = W(t_{i+1}) - W(t_i)$, where $W(t_i)$ is a Wiener process for $i = 1, \ldots, n$.

Itô integral is defined as

$$\int_a^b f(t)\,dt = l.i.m._{n\to\infty} \sum_{k=1}^{n} f(t_k)[W(t_{k+1}) - W(t_k)].$$

The mean square convergence is defined as $l.i.m._{n\to\infty} g(n) = G$ if $\lim_{n\to\infty} E((g(n) - G)^2) = 0$. Note that the Ia function is evaluated at the left hand end point of the subinterval.

- Note that the definition resembles that of Riemann integral.
- But the value of the integral depends on the location of function evaluation.
- Evaluation at other points lead to different rules.
- Convergence is in the mean square sense.

As with Riemann integrals, some simple Itô integrals can be calculated directly using the definition. For example,

$$\int_a^b dW(t) = W(b) - W(a).$$

Properties of Itô integral:
Let $f(t)$ and $g(t)$ be functions with Itô integral defined and α, a, b and c constants such that $a < c < b$, then

- $\int_a^b \alpha f(t)\, dW(t) = \alpha \int_a^b f(t)\, dW(t)$
- $\int_a^b (f(t) + g(t))\, dW(t) = \int_a^b f(t)\, dW(t) + \int_a^b g(t)\, dW(t)$
- $\int_a^b f(t)\, dW(t) = \int_a^c f(t)\, dW(t) + \int_c^b f(t)\, dW(t)$
- $E\left(\int_a^b f(t)\, dW(t)\right) = 0$
- $E\left(\left(\int_a^b f(t)\, dW(t)\right)^2\right) = \int_a^b E\left(f^2(t)\right)\, dt$

Examples

- $\int_a^b dW(t) = l.i.m._{n\to\infty} \sum_{k=1}^{n})[W(t_{k+1}) - W(t_k)] = W(b) - W(a)$
- If $F(W(t), t)$ is Itô integrable then $\int_a^b dF(W(t), t) = F(W(b), b) - F(W(a), a)$
- $\int_0^t W(s)dW(s) = \frac{1}{2}\left(W^2(t) - t\right)$
- $\int_a^b W(t)dW(t) = \frac{1}{2}\left(W^2(b) - W^2(a)\right) - \frac{1}{2}(b-a)$

We investigate the expectation of Itô integrals:

$$
\begin{aligned}
E\int_0^t f(s)dW_s =& E\sum_{k=1}^{n} f(t_k)[W_{k+1}-W_k] \\
&= \sum_{k=1}^{n} E\{f(t_k)[W_{k+1}-W_k]\} \\
&= \sum_{k=1}^{n} E(f(t_k)E[W_{k+1}-W_k]) = 0
\end{aligned}
$$

since each $E(W_k) = 0$.

The variance of Itô integrals leads to the Itô isometry property:

$$E\big(\big(\int_0^t f(s,\omega)dW_s\big)^2\big) = \int_0^t E\big(f^2(s)\big)ds.$$

The Itô calculus version of the chain rule, known as the Itô formula, for $Y_t(\omega) = U(t, X_t(\omega))$ with $dX_t = f(t,\omega)dW_t$, is given by

$$dY_t = \Big(\frac{\partial U}{\partial t} + \frac{1}{2} f^2 \frac{\partial^2 U}{\partial X^2}\Big) dt + f \frac{\partial U}{\partial X} dW_t.$$

Note the extra second order term.

A consequence of the Itô formula is

$$\int_0^t W_s dW_s = W_t^2/2 - t/2.$$

Another consequence of the Itô formula is the Itô-Taylor series, useful in developing numerical schemes.
Let X_t be a stochastic process satisfying

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s) ds + \int_{t_0}^{t} b(s, X_s) dW_s$$

for $t \in [t_0, T]$ for sufficiently smooth a and b. For twice differentiable function $f(t, X_t)$, the Itô formula states that

$$
\begin{aligned}
f(t, X_t) = f(t_0, X_{t_0}) + \int_{t_0}^{t} \frac{\partial f}{\partial t} ds \\
+ \int_{t_0}^{t} \left(a(s, X_s) \frac{\partial f}{\partial X} + \frac{1}{2} b^2(s, X_s) \frac{\partial^2 f}{\partial X^2} \right) ds \\
+ \int_{t_0}^{t} b(s, X_s) \frac{\partial f}{\partial X} dW_s.
\end{aligned}
$$

We apply the Itô formula to a and b in

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)ds + \int_{t_0}^{t} b(s, X_s)dW_s$$

to obtain

$$X_t = X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s + R,$$

where the higher order terms denoted by R are

$$
\begin{aligned}
R = {} & \int_{t_0}^{t} \Big(\int_{t_0}^{s} \big(\frac{\partial a}{\partial t} + L_0 a\big) dz + \int_{t_0}^{s} L_1 a \, dW_z \Big) ds \\
& + \int_{t_0}^{t} \Big(\int_{t_0}^{s} \big(\frac{\partial b}{\partial t} + L_0 \, b\big) dz + \int_{t_0}^{s} L_1 b \, dW_z \Big) dW_s
\end{aligned}
$$

with $L_0 = a\frac{\partial}{\partial X} + \frac{1}{2}b^2\frac{\partial^2}{\partial X^2}$ and $L_1 = b\frac{\partial}{\partial X}$.
This is the first order Itô-Taylor expansion. One obtains the second order expansion by using the Itô formula for $L_1 b$ in the remainder R.

Itô Stochastic Differential Equations

A stochastic process $\{X(t) : t \in [0, \infty)\}$ is said to satisfy an Itô stochastic differential equation (SDE), written as

$$dX(t) = \alpha(X(t), t)dt + \beta(X(t), t)dW(t),$$

if for $t \geq 0$ it is a solution of the corresponding integral equation

$$X(t) = X(0) + \int_0^t \alpha(X(\tau), \tau)\, d\tau + \int_0^t \beta(X(\tau), \tau)\, dW(\tau),$$

where the first integral is a Riemann integral and the second is an Itô stochastic integral.

A more general form of the chain rule for Itô stochastic calculus, known as Itô's formula, is:
If $X(t)$ is a solution of $\alpha(X(t),t)dt + \beta(X(t),t)dW(t)$ and $F(x,t): \mathbb{R} \times [a,b]$ with continuous derivatives $\frac{\partial F}{\partial t}, \frac{\partial F}{\partial x}$ y $\frac{\partial^2 F}{\partial x^2}$, then

$$dF(X(t),t) = f((X(t),t)dt + g(X(t),t)dW(t),$$

con

$$f(x,t) = \frac{\partial F}{\partial t} + \alpha(x,t)\frac{\partial F}{\partial x} + \frac{1}{2}\beta^2(x,t)\frac{\partial^2 F}{\partial x^2}$$

$$g(x,t) = \beta(x,t)\frac{\partial F}{\partial x}.$$

The proof is based on the fact that $\Delta W(t)$ is $O(\sqrt{\Delta t})$.

Itô's formula in integral version is:
Let $X(t)$ be a stochastic process satisfying

$$X(t) = X(t_0) + \int_{t_0}^{t} \alpha(X(s), s) ds + \int_{t_0}^{t} \beta(X(s), s) dW(s)$$

for $t \in [t_0, T]$ and for α y β sufficiently smooth. If $F(X(t), t)$ is a twice differentiable function, the Itô's formula states

$$\begin{aligned} F(X(t), t) = F(X(t_0), t_0) &+ \int_{t_0}^{t} \frac{\partial F}{\partial t} ds \\ &+ \int_{t_0}^{t} \big(\alpha(X(s), s)\frac{\partial F}{\partial x} + \frac{1}{2}\beta^2(X(s), s)\frac{\partial^2 F}{\partial x^2}\big) ds \\ &+ \int_{t_0}^{t} \beta(X(s), s)\frac{\partial F}{\partial x} dW_s. \end{aligned}$$

Example:
To calculate $\int_a^b W(t)\,dW(t)$, consider $X(t) = W(t), F(x,t) = x^2$, then $X(t)$ satisfies $dX(t) = dW(t)$ so that $\alpha = 0$ and $\beta = 1$

$$f(x,t) = \frac{\partial F}{\partial t} + \alpha(x,t)\frac{\partial F}{\partial x} + \frac{1}{2}\beta^2(x,t)\frac{\partial^2 F}{\partial x^2} = 1$$

$$g(x,t) = \beta(x,t)\frac{\partial F}{\partial x} = 2x.$$

Then

$$dF(X(t),t) = f((X(t),t)dt + g(X(t),t)dW(t)$$

gives

$$d\left(W^2(t)\right) = dt + 2W(t)dW(t).$$

Integrating

$$W^2(b) - W^2(a) = b - a + 2\int_a^b W(t)\,dW(t).$$

Example:
Let $X(t) = W(t), F(x,t) = xt$, then $X(t)$ satisfies $dX(t) = dW(t)$ so that $\alpha = 0$ y $\beta = 1$ and Itô's formula gives

$$dF(W(t),t) = d(tW(t)) = \left(\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\right) dt + \frac{\partial F}{\partial x} dW(t)$$

$$= W(t)dt + tdW(t).$$

Integrating

$$bW(b) - aW(a) = \int_a^b W(t)\,dt + \int_a^b t\,dW(t)$$

There is a theorem of existence and uniqueness of solutions of stochastic differential equations ([8])
Theorem: Let $T > 0, \alpha(x,t), \beta(x,t)$ and constants $C > 0, D > 0$ such that

1. $|\alpha(x,t)| + |\beta(x,t)| \leq C(1+|x|)$ for all $x \in \mathbb{R}, t \in [0,T]$
2. $|\alpha(x,t) - \alpha(y(t)| + |\beta(x,t) - \beta(y,t)| \leq D|x-y|$ for all $x, y \in \mathbb{R}, t \in [0,T]$

then

$$dX(t) = \alpha(X(t),t)dt + \beta(X(t),t)dW(t), \quad X(0) = X_0$$

has a unique solution.

Population Models

Example:
Consider Malthus population growth model $dx = ax(t)dt$ and consider that the growth coefficient has a deterministic part and a random part $adt = rdt + hdW(t)$. The stochastic equation is

$$dX(t) = rX(t)dt + hX(t)dW(t),$$

so $\alpha(x,t) = rx$ y $\beta(x,t) = hx$. Itô's formula is

$$dF(= \left(\frac{\partial F}{\partial t} + \alpha\frac{\partial F}{\partial x} + \frac{1}{2}\beta^2\frac{\partial^2 F}{\partial x^2}\right)dt + \beta\frac{\partial F}{\partial x}dW(t).$$

The noise term is called the diffusion term. Noise proportional to the size of the population is known as multiplicative noise.

If we divide by $X(t)$ and use that $d\ln(X(t) = dX(t)/X(t)$ we can take $F(x,t) = \ln(x)$:

$$d\left(\ln(X(t)\right) = \left(rX(t)\frac{1}{X(t)} + \frac{1}{2}h^2X(t)^2\left(-\frac{1}{X(t)^2}\right)\right) + hX(t)\frac{1}{X(t)}dW(t)$$

$$= (r - \frac{h^2}{2})dt + hW(t).$$

Integrating from 0 to t gives

$$\ln\left(\frac{X(t)}{X(0)}\right) = (r - \frac{h^2}{2}t = hW(t)$$

and therefore

$$X(t) = X(0)\exp\left((r - \frac{h^2}{2})t + rW(t)\right).$$

If we integrate the stochastic differential equation and we take the expected value

$$E(X(t)) = E\left(\int_0^t rX(s)\,ds\right) + hE\left(\int_0^t X(s)\,dW(s)\right)$$

but since

$$E\left(\int_a^b f(t)\,dW(t)\right) = 0$$

we have that $\quad E(X(t)) = r\int_0^t E(X(s))\,ds$

or

$$\frac{E(X(t))}{dt} = rE(X(t))$$

with solution

$$E(X(t)) = E(x(0))\exp(rt) = X(0)\exp(rt).$$

Another alternative is to add a constant multiple of dWt to the deterministic equation (additive noise)

$$dX(t) = rX(t)dt + hdW(t),$$

where we are now assuming that the noise is an environmental noise independent of the size of the population.

Using Itô's formula the exact solution is

$$Xt = e^{at} X0 + \int_0^t e^{a(t-s)} b \, dWs.$$

Which of the two formulations is better?
For small populations the environmental noise may be dominant. But since environmental fluctuations often are related to overpopulation in ways such as shortage of food, increased aggression toward each other, etc., a reasonable way to modify the deterministic equation is to consider introducing randomness that is proportional to the size of the population.

Example:
In the logistic population growth model it is assumed that the population growth is limited by the competition for resources between members of the same species and that this competition is proportional to the number of contacts between members of the population, x^2. The deterministic equation is

$$dx = rx\,(1 - x/K)\,dt$$

with K a constant known as the carrying capacity. There are several ways of obtaining a stochastic equation but the simplest one is to assume that the noise is proportional to the size of the population

$$dX(t) = rX(t)\,(1 - x/K)\,dt + \sigma X(t) dW(t).$$

The exact solution is ([2])

$$X(t) = \frac{\exp\left((r - \frac{\sigma^2}{2} t + \sigma W(t)\right)}{X^{-1}(0) + \frac{r}{K}\int_0^t \exp\left((r - \frac{\sigma^2}{2})s + \sigma W(s)\right) ds}.$$

Other equations with exact solution are

- $$dX(t) = \alpha dt + \beta dW(t), \quad \alpha, \beta \in \mathbb{R}$$
  $$X(t) = \alpha t + \beta W(t) + X(0)$$
  .
- Ornstein-Uhlenbeck (it is the same as Malthus equation previously done)
  $$dX(t) = -\alpha X(t)dt + \beta dW(t)$$
  $$X(t) = X(0)\exp(-\alpha t) + \exp(-\alpha t)\int_0^t \exp(\alpha s)\beta dW(s)$$
- $$dX(t) = \alpha(t)X(t)dt + \beta(t)X(t)dW(t)$$
  $$X(t) = X(0)\exp\left(\int_0^t (\alpha(s) - \frac{1}{2}\beta^2(s))\, ds + \int_0^t \beta(s)\, dW(s)\right)$$

Monod growth model

$$\begin{cases} dB(t) = \left(\dfrac{rB(t)\,C(t)}{k_2 + C(t)} - \mu B(t) \right) dt \\[2ex] dC(t) = -\dfrac{k_1 B(t)\,C(t)}{k_2 + C(t)} dt \end{cases}$$

$B(t)$: concentration of bacteria at time t
$C(t)$: concentration nutrients at time t

A system of stochastic differential equations based on Monod's model is

$$\begin{cases} dB(t) = \left(\dfrac{rB(t)C(t)}{k_2 + C(t)} - \mu B(t) \right) dt + eB(t)\, dW(t), \\[2ex] dC(t) = -\dfrac{k_1 B(t) C(t)}{k_2 + C(t)} dt + fC(t)\, dU(t). \end{cases}$$

This system does not have an exact solution.

SIR epidemic model
The deterministic model is

$$\frac{dS}{dt} = -\frac{\beta}{N} SI$$
$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I,$$

where S is the susceptible individuals, I the infective and infectious individuals, and R the recovered with immunity individuals. β is the contact rate, γ the recovery rate and N the total population. Only the first two equations need to be considered.

There are many different ways of obtaining a stochastic system

$$dS = -\frac{\beta}{N} SIdt + b_1 dW_1(t)$$
$$dI = (\frac{\beta}{N} SI - \gamma I)dt + b_2 dW_2(t).$$

The noise is considered environmental

A second way is to consider that the noise depends on the size of the corresponding population

$$
\begin{aligned}
dS &= -\frac{\beta}{N} SIdt + b_1 SdW_1(t) \\
dI &= (\frac{\beta}{N} SI - \gamma I)dt + b_2 IdW_2(t).
\end{aligned}
$$

A third way of obtaining a stochastic system is to consider that some of the coefficients have a deterministic and an a random part

$$
\begin{aligned}
dS &= -\frac{\beta}{N} SIdt + b_1 \frac{\beta}{N} SIdW_1(t) - \gamma I dW_2(t) \\
dI &= (\frac{\beta}{N} SI - \gamma I)dt + b_1 \frac{\beta}{N} SIdW_1(t) + b_2 - \gamma I dW_2(t).
\end{aligned}
$$

A fourth way is to start with a discrete time Markov chain, calculate the probabilities of a population going or exiting a given state. The expectation and covariance matrix are calculatd and under certain assumptions (Central limit theorem and others) a stochastic system is derived. It is not unique ([1])

$$
\begin{aligned}
dS &= -\frac{\beta}{N} SIdt + G_{11} dW_1(t) + G_{12} dW_2(t) \\
dI &= (\frac{\beta}{N} SI - \gamma I)dt + G_{21} dW_1(t) + G_{22} dW_2(t),
\end{aligned}
$$

where

$$
G^2 = \begin{pmatrix} \beta SI/N & -\beta SI/N \\ -\beta SI/N & \beta SI/N + \gamma I \end{pmatrix}
$$

SIRS epidemic model

The deterministic model is

$$\frac{dS}{dt} = -\frac{\beta}{N} SI + cR$$
$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I - cR,$$

where S is the susceptible individuals, I the infective and infectious individuals, and R the recovered with temporary immunity individuals. β is the contact rate, γ the recovery rate and N the total population. Since the total population is constant, R can be eliminated and only the first two equations need to be considered. Stochastic models can be derived similarly to the SIR.

SIRS epidemic model with vaccination Vaccinated individuals get partial immunity and may become susceptible or infected. A fraction α of the newborns are vaccinated and the birth rate μ is equal to the death rate of all compartments, so the total population N is constant.

The deterministic model is

$$\begin{aligned}
\frac{dS}{dt} &= (1-\alpha)\mu N - \frac{\beta}{N}SI - \mu S - \phi S + \delta V + cR \\
\frac{dI}{dt} &= \frac{\beta}{N}SI + \sigma\frac{\beta}{N}VI - (\gamma+\mu)I \\
\frac{dR}{dt} &= \gamma I - (\mu + c)R \\
\frac{dV}{dt} &= \alpha\mu N + \phi S - (\delta+\mu)V - \sigma\frac{\beta}{N}VI,
\end{aligned}$$

where V is the vaccinated individuals, ϕ is the vaccination rate of susceptible individuals, σ the effectiveness of the vaccine (between 0 and 1), and δ the loss of immunity due to the vaccine. Since $V = N - S - I - R$ the last equation can be eliminated.

Numerical methods

A numerical solution of the stochastic differential equation

$$dX(t) = \alpha(X(t), t)dt + \beta(X(t), t)dW(t) \quad \text{in} \quad [0, T], \quad X(0) = X_0$$

is a stochastic process that solves the equivalent equation when the differentials are replaced by difference approximations. Take a partition of $[0, T]$;

$$0 = t_0 < t_1 < \ldots < t_{k-1} < t_k = T.$$

The mesh of the partition is

$$\delta(k) = \max_{i=1,\ldots,k} (t_i - t_{i-1}).$$

A measure of the error of the approximation for a given trajectory is

$$\varepsilon_s(\delta(k)) = E|X(T,\omega) - X_k(T,\omega)|,$$

where $X(T,\omega)$ is the exact solution on a given Brownian path and $X_k(T,\omega)$ is the solution of the difference equation. A strong numerical solution of an Itô stochastic differential equation is a stochastic process $X_k(t)$ that satisfies

$$\varepsilon_s(\delta(k)) \to 0 \text{ as } \delta(k) \to 0.$$

A weak numerical solution only aims at approximating the moments of the solution X. That is its probability distribution function approximates that of the exact solution. The approximation is not path-wise. Let

$$\varepsilon_w(\delta(k)) = E|f(X(T)) - f(X_k(T))|,$$

where the f is usually a polynomial so we get the moments. A weak numerical solution of an Itô stochastic differential equation is a stochastic process $X_k(t)$ that satisfies

$$\varepsilon_w(\delta(k)) \to 0 \text{ as } \delta(k) \to 0.$$

The numerical solution X_k converges strongly with order $\gamma > 0$ if there exists a constant $c > 0$ such that

$$\varepsilon_s(\delta)k)) < c\delta(k)^{\gamma}$$

for all partitions.
The numerical solution X_k converges weakly with order $\gamma > 0$ if there exists a constant $c > 0$ such that

$$\varepsilon_w(\delta)k)) < c\delta(k)^{\gamma}$$

for all partitions.

Euler-Murayama method

Consider the stochastic differential equation

$$dX(t) = \alpha(X(t),t)dt + \beta(X(t),t)dW(t) \quad \text{in} \quad [0,T], \quad X(0) = X_0$$

Take a uniform partition of $[0,T]$;

$$0 = t_0 < t_1 < \ldots < t_{k-1} < t_k = T, \quad \Delta t = t_{i+1} - t_i = \frac{T}{k}$$

So that $t_{i+1} = t_i + \Delta t = i\Delta t$.

$$\text{Y} \quad \Delta W_i = \Delta W(t_i) = W(t_i + \Delta t) - W(t_i).$$

In the method of Euler-Murayama, which is the simplest one,

$$dX(t_i) \approx \Delta X(t_i) = X(t_{i+1}) - X(t_i) \quad \text{y} \quad dW(t_i) \approx \Delta W_i.$$

Therefore

$$X_{i+1} = X_i + \alpha(X_i, t_i)\Delta t + \beta(X_i, t_i)\Delta W_i$$

Since $\Delta W_i \sim N(0,\Delta t)$ if $\eta \sim N(0,1)$ then $\sqrt{\Delta t}\eta \sim N(0,\Delta t)$. To generate a trajectory or realization it is necessary to generate random values of η_i.

$$X_{i+1} = X_i + \alpha(X_i,t_i)\Delta t + \beta(X_i,t_i)\sqrt{\Delta t}\eta_i, \quad i = 0,1,\ldots,k-1$$

$$X(0) = X_0.$$

The error satisfies:

$$E(|X(t_i) - X_i|^2 | X(t_{i-1}) - X_{i-1}) = O(\Delta t^2) \quad \text{(one step)}$$

$$E(|X(t_i) - X_i|^2 | X(0) - X_0) = O(\Delta t).$$

Order of weak convergence equal to 1. And

$$X(t_i) - X_i = O(\Delta t)^{.5}$$

Order of strong convergence equal to .5

Milstein method

The method of Milstein is

$$X_{i+1} = X_i + \alpha(X_i, t_i)\Delta t + \beta(X_i, t_i)\Delta W_i$$
$$+\frac{1}{2}\beta(X_i, t_i)\frac{\partial \beta(X_i, t_i)}{\partial x}\left[(\Delta W_i)^2 - \Delta t\right].$$

The order of strong convergence is 1.

Programs

The next program in octave and Matlab solves Malthus population growth model:

$$dX(t) = r\,X(t)dt + c\,X(t)dW(t).$$

The exact solution is :

$$X(t) = X_0 \exp((r - \frac{1}{2}c^2)t + c\,W(t)).$$

```
clear; clf;
T=1; % maximum time
r=1/2;
c=.1;% coefficients of the  equation
k=100; % number of subintervals
deltat=T/k;
deltat2=sqrt(deltat);
npath=4;% number of trajectories. The last one is deterministic
for j=1:npath
      if j==npath
          c=0.;
     end
      X(1)=1.;
    for i=1:k
         X(i+1)=X(i) +r* X(i)*deltat+ c* X(i)* deltat2*randn;
  end
     plot([0:deltat:T],X, 'LineWidth',2);
```

```
   hold on
end
title('Euler-Murayama');
xlabel('t');
ylabel('X(t)');
```

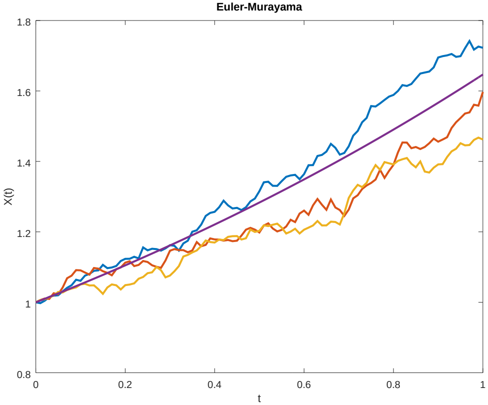

Figure: Method of Euler for the model of Malthus. The smooth trajectory is the deterministic solution.

The next code in octave and Matlab solves the logistic equation using the method of Euler

```
clear; clf;
T=10; %maximum time
r=1/2;
b=.1;% coefficients of the equation
k=100; % number of subintervals
K=2% carrying capacity
deltat=T/k;
deltat2=sqrt(deltat);
npath=4;% number of trajectories; the last one is deterministic
for j=1:npath
    if j==npath
        b=0.;
    end
     X(1)=1.;
   for i=1:k
        X(i+1)=X(i) +r* X(i)*(1.-X(i)/K)*deltat+ b* X(i) * deltat2*randn;
  end
```

```
      plot([0:deltat:T],X, 'LineWidth',2);
   hold on
end
title('Euler-Murayama');
xlabel('t');
ylabel('X(t)');
```

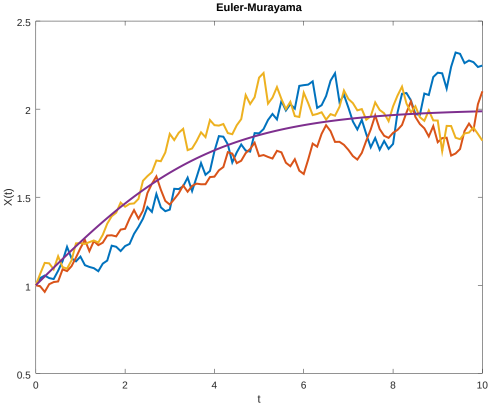

Figure: Euler method for the logistic model. The smooth trajectory is the deterministic solution.

The following program calculates the mean of 500 trajectories produced by the methods of Euler and Milstein for the model of Malthus and compares them with the expected value of the exact solution of the stochastic equation.

```
%Compares the average of Euler and Milstein with the
    exact for the equation  dX=rX dt +cX dW, X(0)=X0
   clear; clf;
nsim=500; % number of trajectories
T=3; % maximum time
r=1/2;
c=.1;% coefficients of the equation
k=30; k1=k+1;% number of subintervals
deltat=T/k;
deltat2=sqrt(deltat);
X0=1.;
Xe=zeros(nsim,k1);
Xm=zeros(nsim,k1);
for m=1:nsim
      Xe(m,1)=X0;
   Xm(m,1)=X0;
for i=1:k
     dw=deltat2*randn;
  Xe(m,i+1)=Xe(m,i)+r*Xe(m,i)*deltat+c*Xe(m,i)*dw;
```

```
    Xm(m,i+1)=Xm(m,i) +r* Xm(m,i)*deltat+ c* Xm(m,i)* dw +...
    .5*c*Xm(m,i)*c*(dw^2-deltat);
end
end
meanXe=sum(Xe)/nsim
meanXm=sum(Xm)/nsim
EX=X0*exp(r*(0:deltat:T));
     plot([0:deltat:T],EX, 'k-',[0:deltat:T],meanXe, 'r--*', [0:deltat:T], meanXm, 'b-.x','LineWidth',2);
title('Euler-Milstein-Exact');
xlabel('t');
ylabel('X(t)');
legend('Expected value','Mean Euler-Murayama','Mean Milstein')
diffXe=abs(meanXe(k)-EX(k))
diffXm=abs(meanXm(k)-EX(k))
```

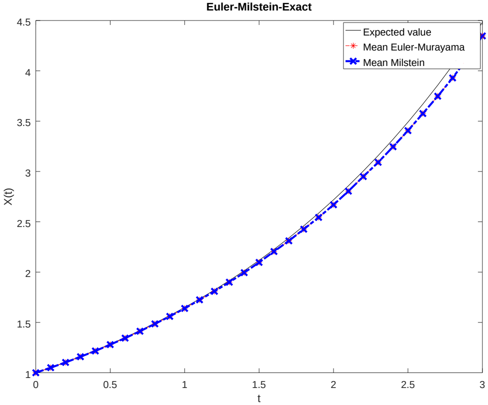

Figure: means of the methods of Euler and Milstein for Malthus model, and the expected value of the exact solution.

The next program solves the Monod model using the method of Milstein

```
%Method of  Milstein for the Monod system and deterministic solution
%dB=(rBC/(k2+C)-mu B)dt + s1 B dW; dC=-k3B C/(k2+C) dt+ s2 C dU
    clear; clf;
T=5; % maximum time
r=2.;
k2=1.; k3=1.; mu=1/2; s1=.05; s2=.05;% coefficients of the equation
k=200; k1=k+1;% number of subintervals
deltat=T/k;
deltat2=sqrt(deltat);
B0=1.; B(1)=B0; Bd(1)=B0;
C0=10.; C(1)=C0; Cd(1)=C0;
randn('state',10);
for i=1:k
    dw=randn(2,1)
```

```
  B(i+1)=B(i)+(r*B(i)*C(i)/(k2+C(i))-mu*B(i))*deltat+s1*B(i)*dw(1)+...
        s1^2*B(i)*(dw(1)^2-deltat)/2.;
  C(i+1)=C(i)-r*B(i)*C(i)/(k2+C(i))*deltat+s2*C(i)*dw(2)+...
        s2^2*C(i)*(dw(2)^2-deltat)/2.;
  Bd(i+1)=Bd(i)+(r*Bd(i)*Cd(i)/(k2+Cd(i))-mu*Bd(i))*deltat;
  Cd(i+1)=Cd(i)-r*Bd(i)*Cd(i)/(k2+Cd(i))*deltat;
end
   plot([0:deltat:T],Bd, 'k-',[0:deltat:T],B, 'r--*', [0:deltat:T], Cd, 'b-.x', [0:deltat:T], C,'g--+','LineWidth',2);
title('Monod Milstein and deterministic');
xlabel('t');
ylabel('B(t), C(t)');
legend('Bacteria deterministic','Bacteria Milstein','Nutrient deterministic', 'Nutrient Milstein')
```

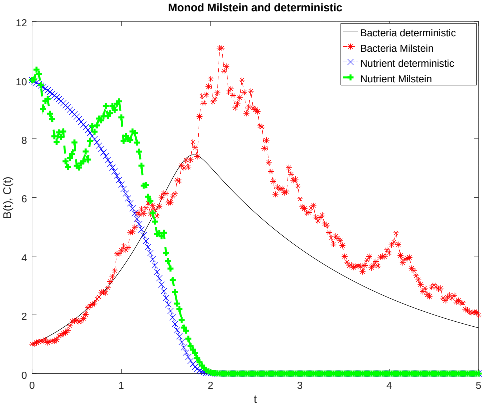

Figure: Method of Milstein for the model of Monod.

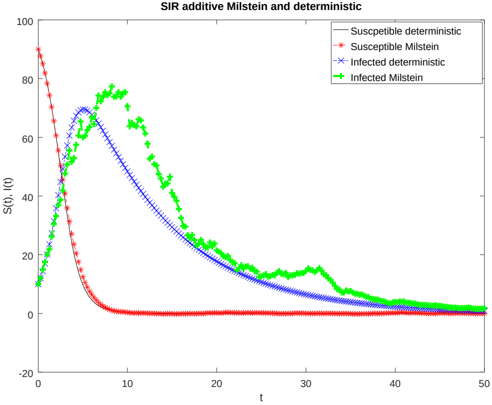

Figure: Method of Milstein for the SIR model with environmental (additive) noise. The smooth trajectory is the deterministic solution.

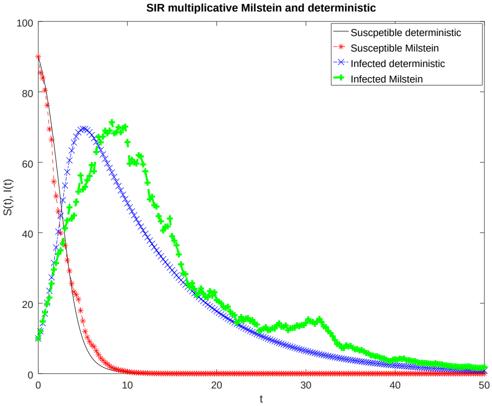

Figure: Method of Milstein for the SIR model with multiplicative noise (proportional to the population size). The smooth trajectory is the deterministic solution.

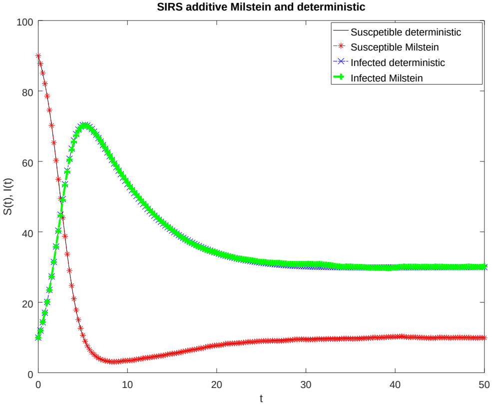

Figure: Method of Milstein for the SIRS model with low additive noise. The smooth trajectory is the deterministic solution.

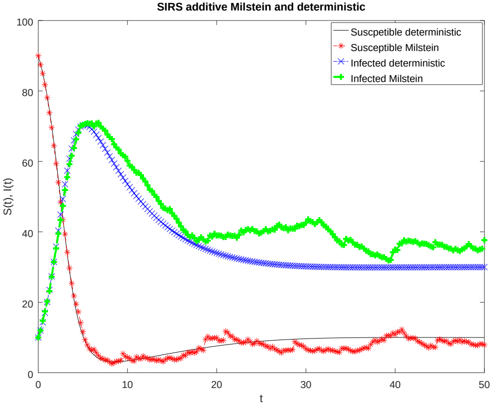

Figure: Method of Milstein for the SIRS model with higher additive noise. The smooth trajectory is the deterministic solution.

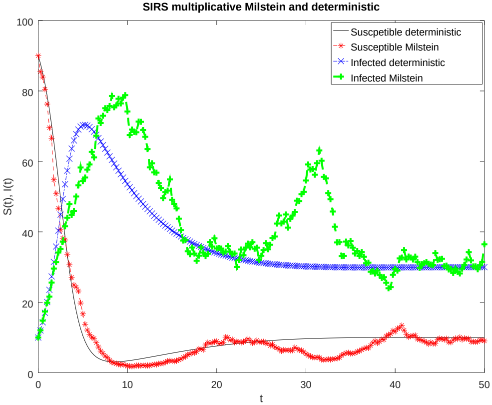

Figure: Method of Milstein for the SIRS model with multiplicative noise. The smooth trajectory is the deterministic solution.

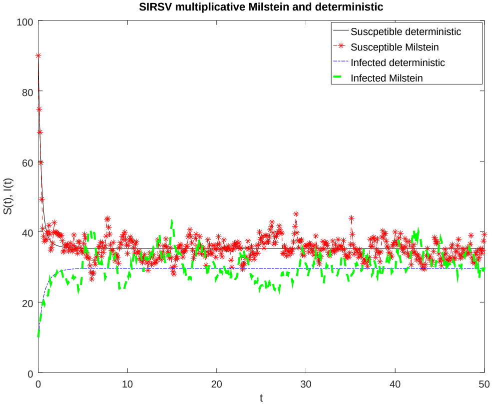

Figure: Method of Milstein for the SIRSV model with multiplicative noise and medium vaccine efficiency. The smooth trajectory is the deterministic solution.

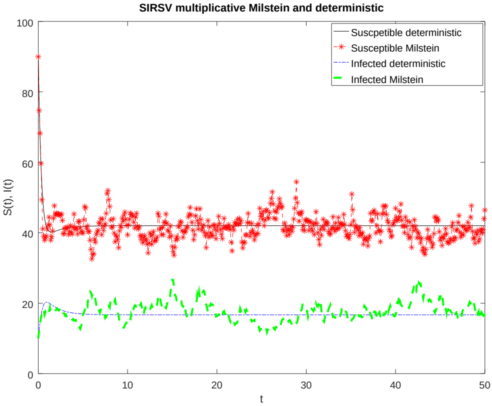

Figure: Method of Milstein for the SIRSV model with multiplicative noise and total efficiency $\sigma = \delta = 0$. The smooth trajectory is the deterministic solution.

- The paper by Higham [4] has programs in octave and Matlab. The programs are in `http://personal.strath.ac.uk/d.j.higham/algfiles.html`
- Software package SDEtoolbox has programs for several models. It can be downloaded from `http://sdetoolbox.sourceforge.net/`

Allen, E. , 2007, *Modeling with Ito Stochastic Differential Equations*, Springer, Dordrecht, The Netherlands.

Allen, L., 2010, *An Introduction to Stochastic Processes with Applications to Biology*, 2nd edition, CRC Press, Boca Raton.

Evans, L. C., *An Introduction to Stochastic Differential Equations*, Version 1.2, http://www.gaianxaos.com/pdf/stochastics/stochastic_diffeq.pdf

Higham, D. J., 2001, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM REVIEW, Society for Industrial and Applied Mathematics Vol. 43, pp. 525–546.

Ito, K., and McKean, H.P.,Jr., 1965, *Diffusion Processes and Their Sample Paths*, Academic Press. N.Y.

Kloeden, P.E., and Platen, E., 1995, *Numerical Solution of Stochastic Differential Equations*, Springer. N.Y.

McKean, H.P.,Jr., 1969, *Stochastic Integrals*, Academic Press. N.Y.

Oksendal, B., 1995, *Stochastic Differential Equations. An Introduction with Applications*, 4th edition. Springer. N.Y.

Schuss, Z., 1980, *Theory and Applications of Stochastic Differential Equations*, John Wiley and Sons. N.Y.