Comparison of Probabilistic and Stochastic Formulations in Modeling Growth Uncertainty and Variability

Shuhua Hu

Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695

April 25, 2008

Acknowledgments

- Center for Research in Scientific Computation, North Carolina State University
  Dr. H. T. Banks, J. L. Davis, and S. L. Ernstberger
- Advanced BioNutrition Corporation (ABN)
  Dr. A. K. Dhar, and E. Artimovich
- Waddell Marine Culture Center: Dr. C. L. Browdy

Reference

- Comparison of Probabilistic and Stochastic Formulations in Modeling Growth Uncertainty and Variability (with H. T. Banks, J. L. Davis, S. L. Ernstberger, A. K. Dhar, C. L. Browdy and E. Artimovich), CRSC-TR08-03, North Carolina State University, 2008.

Outline of the Talk

- Research Motivation
- Difference between Probabilistic Formulation and Stochastic Formulation in Modeling Growth Uncertainty
- Include Comparable Amount of Uncertainty in Probabilistic and Stochastic Formulations to Make Reasonable Comparison
- Conclusion

Research Motivation

- Use shrimp as biofactory for production of biological countermeasures responding to national emergency situations such as bio toxic attacks on population
- Stock shrimp postlavae in a marine raceway and allow them to grow to juvenile stage, and then infect them using a virus carrying a passenger gene for the desired countermeasure.

Raceway for Culturing Shrimp

- **Mathematical Goal**

  Provide mathematical models for the whole production system: biomass production for the healthy shrimp, and vaccine production for the infected shrimp.

  Biomass production sytem: size-structured population model (Sinko and Streifer in 1967)

$$u_t(x,t) + (g(x)u(x,t))_x + m(x)u(x,t) = 0,$$

$$u(0,t) = 0,$$

$$u(x,0) = u_0(x).$$

  In this model, each individual grows according to the *deterministic growth model*

$$\frac{dx}{dt} = g(x) \Rightarrow \text{ individuals with same size have same growth rate.}$$

  No dispersion in size if all individuals start at same size.

- **Experimental Data of Shrimp**

  Fifty shrimp were randomly sampled from each raceway, but are not guaranteed to be sampled from the same set of individuals at different time points (individuals are not tracked over time). Aggregate type longitudinal data

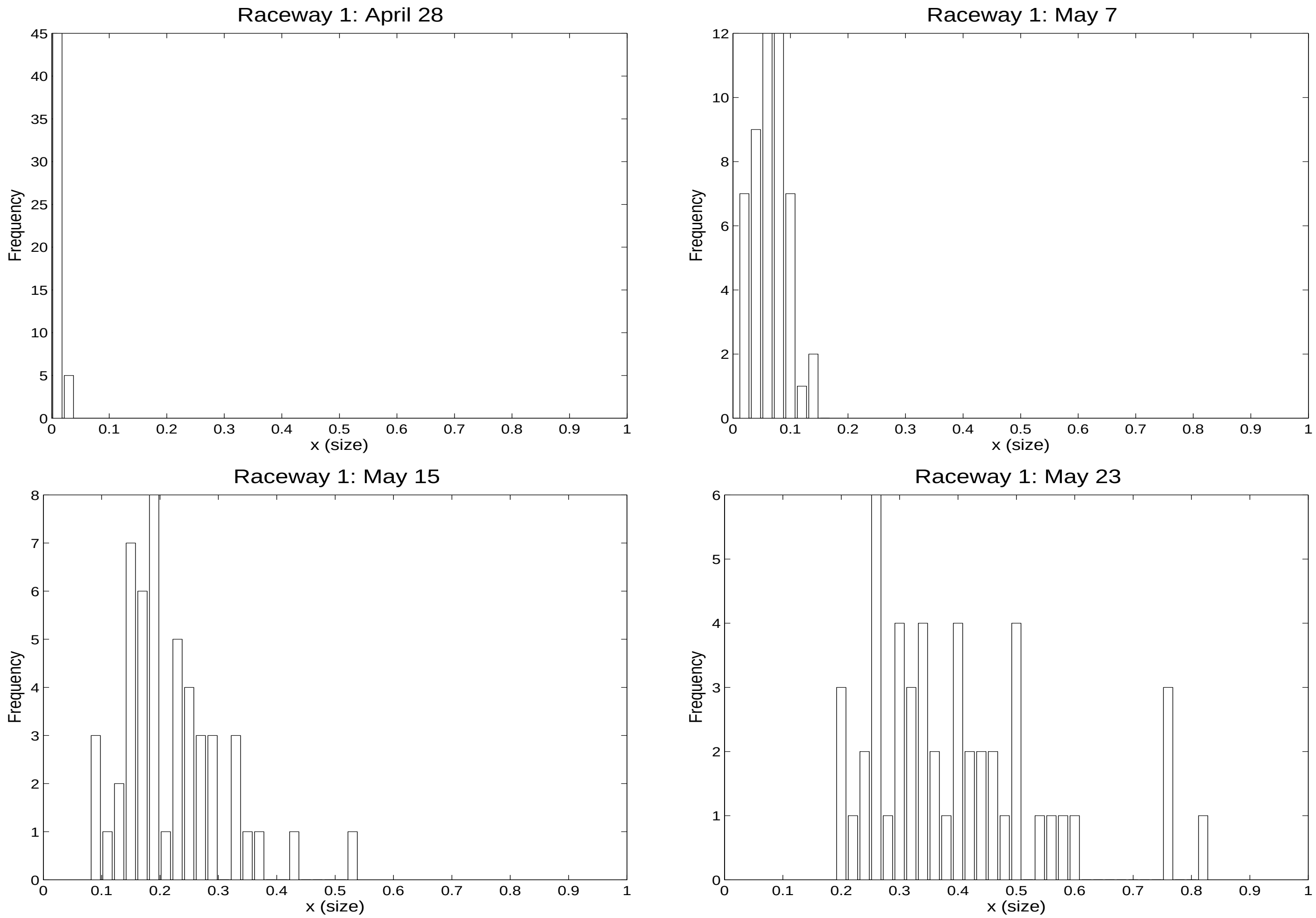

Figure 1: Histograms for longitudinal data for raceway 1.

Remark on Size-Structured Data

- Shrimp population exhibits a great deal of variability in size as time progresses even though they begin at similar size.
- Size-structured population model is **unable** to describe the growth dynamics of this population!
- Need to incorporate some type of variability or uncertainty into the growth of shrimp in addition to simple size dependence so that variability in size distribution is not only determined by the variability in initial size but also the variability or uncertainty in individual growth.
- Two approaches have been used in the literature to model growth uncertainty: probabilistic formulation and stochastic formulation.

Probabilistic Formulation vs. Stochastic Formulation

Probabilistic Formulation

- **Motivation**: effect of genetic differences or some chronic disease on the growth of individuals.
- **Assumption**

  Each individual grows according to a deterministic growth model, but different individuals even with the same size may have different growth rate.

  Partition the entire population into (possible a continuum of) subpopulations where individuals in each subpopulation have the same size-dependent growth rate, and then impose a probability distribution to this partition of possible growth rates in the population.
- **Individual level: deterministic growth**

  The growth process for individuals in a subpopulation with the rate g is described by

$$\frac{dx(t;g)}{dt} = g(x(t;g),t), \;\; g \in \mathcal{G},$$

where $\mathcal{G}$ is a collection of admissible growth rates.

- **Population level: growth rate distribution (GRD) model**

  Each subpopulation is described by a size-structured population model. Let $v(x,t;g)$ be the population density of individuals with size x at time t in a subpopulation with growth rate g.

$$v_t(x,t;g) + (g(x,t)v(x,t;g))_x = 0,$$

$$v(0,t;g) = 0, \;\; v(x,0;g) = v_0(x;g),$$

  for a given $g \in \mathcal{G}$. Then the expected population density with size x at time t is given by

$$u(x,t) = \int_{g\in\mathcal{G}} v(x,t;g)d\mathcal{P}(g),$$

  where $\mathcal{P}$ is a probability measure on $\mathcal{G}$.

- **Remark**

  – This probabilistic structure $\mathcal{P}$ on $\mathcal{G}$ is then the fundamental "parameter" to be determined from aggregate data for the population by using parametric or non-parametric method.

  – Thus this *probabilistic formulation* involves a *stationary probabilistic structure* on a *family of deterministic dynamical systems.*

Stochastic Formulation

- **Motivation**

  Influence of fluctuations of environment on growth rate of individuals, e.g., growth rate of shrimp are affected by temperature, salinity, dissolved oxygen level, PH level, un-ionized ammonia level, etc.

- **Assumption**

  Movement from one size class to another one can be described by a diffusion process.

- **Individual level: stochastic growth**

  Let $\{X(t) : t \geq 0\}$ be a diffusion process and $X(t)$ denote the size of an individual in the population at time t. Then $X(t)$ is described by the following Ito stochastic differential equation

  $$dX(t) = g(X(t))dt + \sigma(X(t), t)dW(t),$$

  where $W(t)$ is the standard Wiener process.

  Each individual grows according to the above stochastic differential equation.

- **Population level: Fokker-Planck (FP) model**

  This assumption on growth process leads to *Fokker-Planck* (FP) or *forward Kolmogorov* model for population density u (carefully derived by A. Okubo)

$$
\begin{aligned}
&u_t(x,t) + (g(x)u(x,t))_x = \frac{1}{2}(\sigma^2(x,t)u(x,t))_{xx},\\
&g(0,t)u(0,t) - \frac{1}{2}(\sigma^2(x,t)u(x,t))_x|_{x=0} = 0,\\
&g(L,t)u(L,t) - \frac{1}{2}(\sigma^2(x,t)u(x,t))_x|_{x=L} = 0,\\
&u(x,0) = u_0(x).
\end{aligned}
$$

- **Remark**

  Note that FP model with $\sigma = 0$ (no uncertainty in individual growth) yields size-structured population model.

Remark on Probabilistic and Stochastic Formulations

- Probabilistic and stochastic formulations are conceptually quite different.
  - Probabilistic formulation results in GRD model, and growth process for each individual is a deterministic one. Growth uncertainty is introduced into population by the variability of growth rates among individuals.
  - Stochastic formulation results in FP model, and growth process for each individual is a stochastic one. Growth uncertainty is introduced into population by the stochastic growth of each individual.

- The choice of a formulation to describe the dynamics of a particular population should, if possible, be based on the type of scenario causing the uncertainty of growth.

Data Fitting of Mean Size of Shrimp

- **Use exponential function $\bar{\mathbf{x}}(\mathbf{t}) = \mathbf{a}\exp(\mathbf{bt}) + \mathbf{c}$ to fit the mean size of shrimp from raceways 1 and 2.**

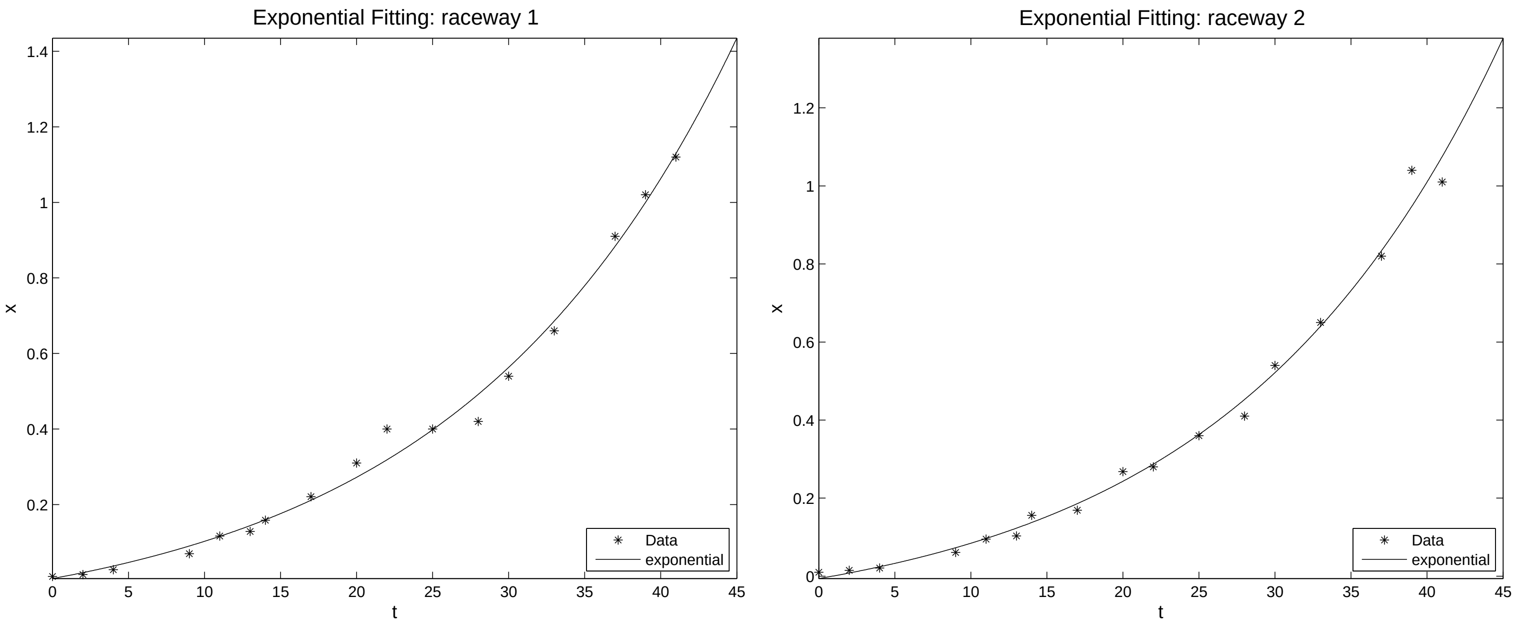

Figure 2: Exponential fit of raceways 1 and 2. (left): $g(\bar{x}) = 0.054(\bar{x}+0.133)$; (right): $g(\bar{x}) = 0.056(\bar{x}+0.126)$.

Remark on Data Fitting

- Results show that $\bar{x}' = g(\bar{x}) = b_0(\bar{x} + c_0)$ is a reasonable description of the early growth of shrimp.

- Let $X(t)$ be a random variable which denotes the size at time t. That is, each realization corresponds to the size of an individual at time t.

Mean growth dynamics: $$\frac{d\mathrm{E}(X(t))}{dt} = b_0(\mathrm{E}(X(t)) + c_0)$$

- Deterministic growth model

$$\frac{dx}{dt} = b_0(x + c_0)$$

Probabilistic formulation

$$\frac{dx}{dt} = b(x + c_0),\ b \in \mathcal{R}_b; \quad \frac{dx}{dt} = b_0(x + c),\ c \in \mathcal{R}_c; \quad \frac{dx}{dt} = b(x + c),\ (b, c) \in \mathcal{R}_b \times \mathcal{R}_c$$

Stochastic formulation

$$dX(t) = b_0(X(t) + c_0)dt + \sigma(x, t)dW(t)$$

Investigations

- Does the **probabilistic formulation** generate a **stochastic process** for size? If so, does the **stochastic process** satisfy a ***random differential equation*** or ***stochastic differential equation***?

Investigations

- Does the **probabilistic formulation** generate a **stochastic process** for size? If so, does the **stochastic process** satisfy a ***random differential equation*** or ***stochastic differential equation***?

  Yes, probabilistic formulation generates a stochastic process, and it satisfies a random differential equation.

Investigations

- Does the **probabilistic formulation** generate a **stochastic process** for size? If so, does the **stochastic process** satisfy a ***random differential equation*** or ***stochastic differential equation***?

IDEA: Put distribution on **b**, **c**, and $\mathbf{x_0}$ simultaneously in $g(x) = b(x+c)$ in the deterministic system

$$\frac{dx(t)}{dt} = b(x+c), \quad x(0) = x_0.$$

More generally, using its solution

$$x(t; b, c, x_0) = (x_0 + c)\exp(bt) - c,$$

and assuming that B, C and X_0 are random variables for b, c and x_0, respectively, we can always define a stochastic process

$$X(t; B, C, X_0) = (X_0 + C)\exp(Bt) - C,$$

and argue that it satisfies the **random differential equation**

$$\frac{dX(t)}{dt} = B(X(t) + C), \quad X(0) = X_0.$$

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on **b** always satisfy the **mean growth dynamics**?

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on $\mathbf{b}$ always satisfy the **mean growth dynamics**?

  NO. For example, if we assume that $B \sim \mathcal{N}(b_0, \sigma_0^2)$, then the expectation of obtained size distribution does not satisfy the mean growth dynamics.

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on **b** always satisfy the **mean growth dynamics**?

  NO. For example, if we assume that $B \sim \mathcal{N}(b_0, \sigma_0^2)$, then the expectation of obtained size distribution does not satisfy the mean growth dynamics.

$$\frac{dx(t)}{dt} = b(x(t) + c_0), \; x(0) = x_0.$$

The size of each individual in a subpopulation with intrinsic growth rate b at time t is

$$x(t; b) = -c_0 + (x_0 + c_0)\exp(bt).$$

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on **b** always satisfy the **mean growth dynamics**?

  NO. For example, if we assume that $B \sim \mathcal{N}(b_0, \sigma_0^2)$, then the expectation of obtained size distribution does not satisfy the mean growth dynamics.

$$\frac{dx(t)}{dt} = b(x(t) + c_0), \; x(0) = x_0.$$

The size of each individual in a subpopulation with intrinsic growth rate b at time t is

$$x(t; b) = -c_0 + (x_0 + c_0)\exp(bt).$$

Let $X(t) = x(t; B)$. Then we have $X(t) = -c_0 + (x_0 + c_0)\exp(Bt)$.

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on **b** always satisfy the **mean growth dynamics**?

NO. For example, if we assume that $B \sim \mathcal{N}(b_0, \sigma_0^2)$, then the expectation of obtained size distribution does not satisfy the mean growth dynamics.

$$\frac{dx(t)}{dt} = b(x(t) + c_0), \ \ x(0) = x_0.$$

The size of each individual in a subpopulation with intrinsic growth rate b at time t is

$$x(t; b) = -c_0 + (x_0 + c_0) \exp(bt).$$

Let $X(t) = x(t; B)$. Then we have $X(t) = -c_0 + (x_0 + c_0) \exp(Bt)$. This implies that $X(t)$ is log-normally distributed for any fixed t, and its mean is given as follows:

$$\mathrm{E}(X(t)) = -c_0 + (x_0 + c_0) \exp\left(b_0 t + \frac{1}{2}\sigma_0^2 t^2\right).$$

Hence, we have

$$\frac{dE(X(t))}{dt} = (b_0 + \sigma_0^2 t)(E(X(t)) + c_0).$$

- Does the **expectation of size distribution** obtained by **probabilistic formulation** with a distribution on **b** always satisfy the **mean growth dynamics**?

  NO. For example, if we assume that $B \sim \mathcal{N}(b_0, \sigma_0^2)$, then the expectation of obtained size distribution does not satisfy the mean growth dynamics.

$$\frac{dx(t)}{dt} = b(x(t) + c_0), \;\; x(0) = x_0.$$

  The size of each individual in a subpopulation with intrinsic growth rate b at time t is

$$x(t; b) = -c_0 + (x_0 + c_0)\exp(bt).$$

  Let $X(t) = x(t; B)$. Then we have $X(t) = -c_0 + (x_0 + c_0)\exp(Bt)$. This implies that $X(t)$ is log-normally distributed for any fixed t, and its mean is given as follows:

$$\mathrm{E}(X(t)) = -c_0 + (x_0 + c_0)\exp\left(b_0 t + \frac{1}{2}\sigma_0^2 t^2\right).$$

  Hence, we have

$$\frac{dE(X(t))}{dt} = (b_0 + \sigma_0^2 t)(E(X(t)) + c_0).$$

  Remark: But the answer is always YES if we put a distribution on **c**.

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

  YES, the expectation of obtained size distribution satisfies the mean growth dynamics.

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

  YES, the expectation of obtained size distribution satisfies the mean growth dynamics.

  Consider stochastic differential equation

$$dX(t) = b_0(X(t) + c_0)dt + \sigma_0(X(t) + c_0)dW(t), \; X(0) = x_0.$$

  By using Ito's formula, we find

$$X(t) = -c_0 + (x_0 + c_0)\exp\left((b_0 - \tfrac{1}{2}\sigma_0^2)t + \bar{\sigma}_0 W(t)\right).$$

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

  YES, the expectation of obtained size distribution satisfies the mean growth dynamics.

  Consider stochastic differential equation

$$dX(t) = b_0(X(t) + c_0)dt + \sigma_0(X(t) + c_0)dW(t), \;\; X(0) = x_0.$$

  By using Ito's formula, we find

$$X(t) = -c_0 + (x_0 + c_0) \exp\left((b_0 - \tfrac{1}{2}\sigma_0^2)t + \bar{\sigma}_0 W(t)\right).$$

  Note that $\mathbf{W}(\mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \mathbf{t})$.

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

  YES, the expectation of obtained size distribution satisfies the mean growth dynamics.

  Consider stochastic differential equation

$$dX(t) = b_0(X(t) + c_0)dt + \sigma_0(X(t) + c_0)dW(t), \; X(0) = x_0.$$

  By using Ito's formula, we find

$$X(t) = -c_0 + (x_0 + c_0) \exp \left((b_0 - \tfrac{1}{2}\sigma_0^2)t + \bar{\sigma}_0 W(t) \right).$$

  Note that $\mathbf{W}(\mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \mathbf{t})$. Hence, $X(t)$ is log-normally distributed for any fixed t, and its mean is given as follows:

$$\mathrm{E}(X(t)) = -c_0 + (x_0 + c_0) \exp(b_0 t).$$

  Hence, we have

$$\frac{dE(X(t))}{dt} = b_0(\mathrm{E}(X(t)) + c_0).$$

- Does the **expectation of size distribution** obtained by **stochastic formulation** with $\sigma(x) = \sigma_0(x + c_0)$ satisfy the **mean growth dynamics**?

  YES, the expectation of obtained size distribution satisfies the mean growth dynamics.

  Consider stochastic differential equation

$$dX(t) = b_0(X(t) + c_0)dt + \sigma_0(X(t) + c_0)dW(t), \; X(0) = x_0.$$

  By using Ito's formula, we find

$$X(t) = -c_0 + (x_0 + c_0)\exp\left((b_0 - \tfrac{1}{2}\sigma_0^2)t + \bar{\sigma}_0 W(t)\right).$$

  Note that $\mathbf{W}(\mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \mathbf{t})$. Hence, $X(t)$ is log-normally distributed for any fixed t, and its mean is given as follows:

$$\mathrm{E}(X(t)) = -c_0 + (x_0 + c_0)\exp(b_0 t).$$

  Hence, we have

$$\frac{dE(X(t))}{dt} = b_0(\mathrm{E}(X(t)) + c_0).$$

  Remark: The answer is also YES for $\sigma(x,t) = \sqrt{2t}\sigma_0(X(t) + c_0)$.

- Can **probabilistic formulation** yield the **same size distribution** as **stochastic formulation** with appropriately chosen parameters?

Yes. Can argue that the size distribution obtained from the stochastic formulation is exactly the same as that obtained from probabilistic formulation if we consider the models:

Stochastic formulation:

$$dX(t) = b_0(X(t) + c_0)dt + \sqrt{2t}\sigma_0(X(t) + c_0)dW(t)$$

Probabilistic formulation:

$$\frac{dx(t;b)}{dt} = (b - \sigma_0^2 t)(x(t;b) + c_0), b \in \mathbb{R} \text{ with } B \sim \mathcal{N}(b_0, \sigma_0^2),$$

with their initial size distributions $X(0)$ the same (either deterministic or random).

Numerical Results: numerical solutions to FP model and GRD model

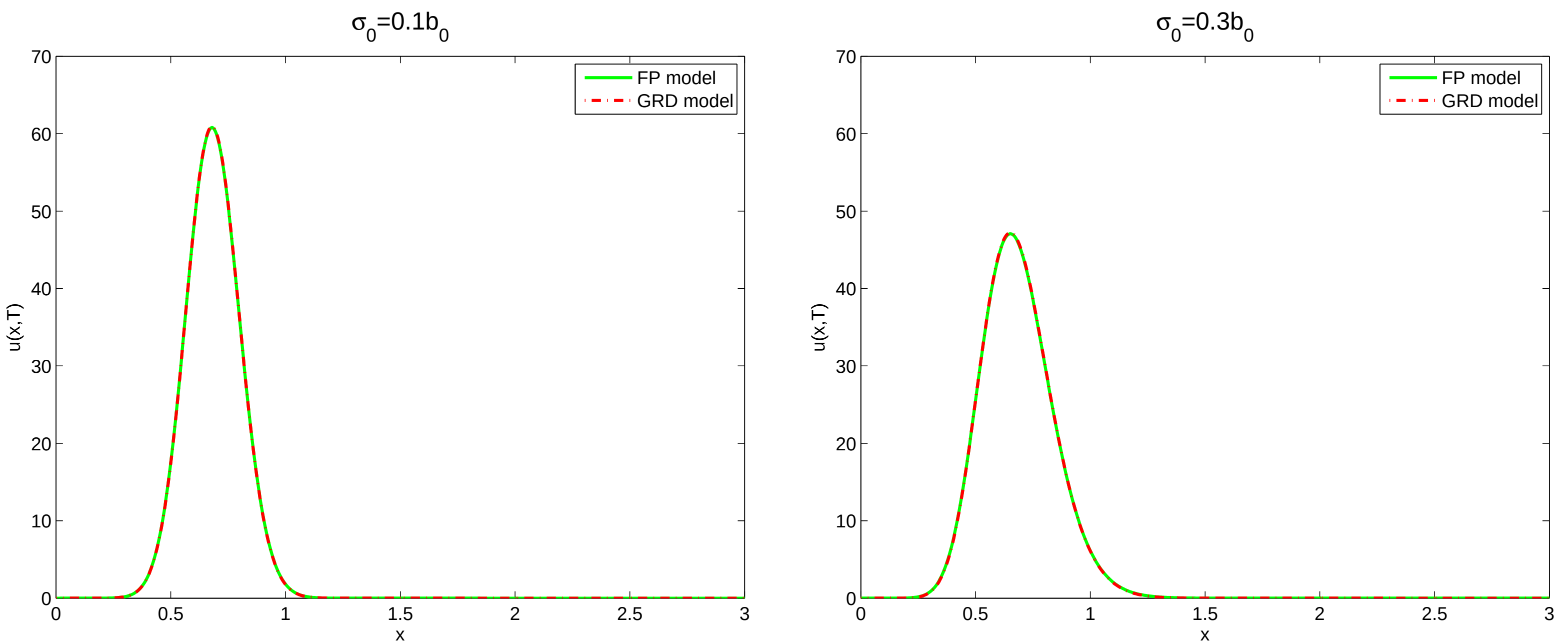

Figure 3: Population density $u(x, T)$ with $T = 10$, $c_0 = 0.1$, $b_0 = 0.045$, and $\sigma_0 = rb_0$. In FP model: $u_0(x) = 100\exp(-100(x-0.4)^2)$. In GRD model: $v_0(x; b) = 100\exp(-100(x-0.4)^2)$ in GRD model, where $B \sim \mathcal{N}_{[\underline{b},\bar{b}]}(b_0, \sigma_0^2)$, $b \in [\underline{b}, \bar{b}]$, $\underline{b} = b_0 - 3\sigma_0$, $\bar{b} = b_0 + 3\sigma_0$. (left): $r = 0.1$; (right): $r = 0.3$.

Conclusion

- The growth process in probabilistic formulation is deterministic in the context of a probability structure, and the growth process in stochastic formulation is stochastic.

- The stochastic process generated by probabilistic formulation satisfies *random differential equation*, but the stochastic process obtained by stochastic formulation satisfies *stochastic differential equation*.

- The expectation of size distribution obtained from both stochastic formulation and probabilistic formulation with distribution on the affine growth satisfy mean growth dynamics, but this is not true for probabilistic formulation with a distribution on the intrinsic growth rate.

- Even though probabilistic and stochastic formulations are conceptually quite different, they can yield the same size distribution with properly chosen parameters.