

Computational Methods for Random Epidemiological Models

Métodos Computacionales para el Estudio de Modelos Epidemiológicos con Incertidumbre

Conferencias de Investigación para Posgrado 2016
Universidad Complutense de Madrid

24 junio de 2016

Prof. Dr. Juan Carlos Cortés

Instituto Universitario de Matemática Multidisciplinar
Universitat Politècnica de València

*im*²



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Part I

Ingredients

A naive (but maybe useful) comparison:

Deterministic	Random
numbers: $a = 3$	r.v.'s: $A \sim N(\mu = 3; \sigma^2 > 0)$
functions: $x(t) = 3t$	s.p.'s: $X(t) = At$, $A \sim N(\mu = 3; \sigma^2 > 0)$

There are s.p.'s which are not defined by algebraic formulas as

Wiener process or Brownian motion

$\{W(t) : t \geq 0\} \equiv \{B(t) : t \geq 0\}$ is called the (standard) **Wiener process** or **Brownian motion** if it satisfies the following conditions:

- ① **It starts at zero w.p. 1:** $\mathbb{P}[\{\omega \in \Omega : W(0)(\omega) = 0\}] = \mathbb{P}[W(0) = 0] = 1$.
- ② **It has stationary increments:**

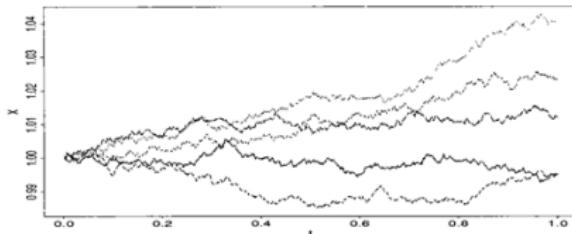
$$W(t) - W(s) \stackrel{d}{=} W(t+h) - W(s+h), \quad \forall h: s, t, s+h, t+h \in [0, +\infty[.$$

- ③ **It has independent increments:**

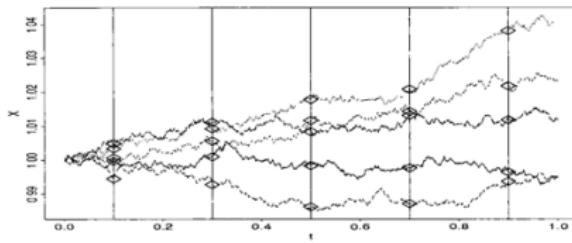
$W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1})$ are independent r.v.'s
 $\forall \{t_i\}_{i=1}^n : 0 \leq t_1 < t_2 < \dots < t_{n-1} < t_n < +\infty$, $n \geq 1$.

- ④ **It is Gaussian with mean zero and variance t :** $W(t) \sim N(0; t)$, $\forall t \geq 0$.

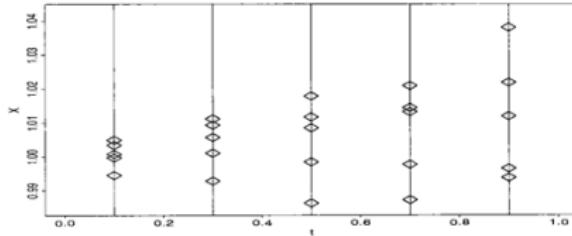
Graphical representation of a s.p.



← trajectories



← trajectories + r.v.'s



← r.v.'s

Since a s.p. $X(t) = \{X(t) : t \in \mathcal{T}\}$ can be considered as a collection of random vectors $(X_{t_1}, \dots, X_{t_n})$, $t_1, \dots, t_n \in \mathcal{T}$, $n \geq 1$, we can extend the concept of expectation and covariance for random vectors to s.p.'s and consider these quantities as functions of $t \in \mathcal{T}$:

One-dimensional probabilistic description of a s.p.

Expectation, variance and 1-p.d.f. of a s.p.

- **Expectation:** $\mu_X(t) = \mathbb{E}[X(t)]$, $t \in \mathcal{T}$.
- **Variance:** $\sigma_X^2(t) = \mathbb{V}[X(t)] = \mathbb{E}[(X(t))^2] - (\mathbb{E}[X(t)])^2$, $t \in \mathcal{T}$.
- **1-p.d.f.:** It is the p.d.f. of the r.v. $X(t)$ for every t . It is denoted by $f_1(x, t)$.

Two-dimensional probabilistic description of a s.p.

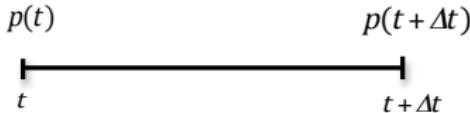
Covariance and 2-p.d.f. of a s.p.

- **Covariance:** $C_X(t_1, t_2) = \mathbb{C}[X_{t_1}, X_{t_2}] = \mathbb{E}[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$, $t_1, t_2 \in \mathcal{T}$.
- **2-p.d.f.:** It is the joint p.d.f. of the r.v.'s $X(t_1)$ and $X(t_2)$ for every t_1 and t_2 . It is denoted by $f_2(x_1, t_1; x_2, t_2)$.

Part II

Linear Models

Motivating the linear case: The malthusian population model with migration



$$p(t + \Delta t) - p(t) = \underbrace{bp(t)\Delta t}_{\text{births}} - \underbrace{dp(t)\Delta t}_{\text{deaths}} + \underbrace{i\Delta t}_{\text{immigrants}} - \underbrace{e\Delta t}_{\text{emigrants}},$$

$$p(t + \Delta t) - p(t) = kp(t)\Delta t + m\Delta t, \quad k = b - d, \quad m = i - e \in \mathbb{R},$$

$$\frac{p(t + \Delta t) - p(t)}{\Delta t} = kp(t) + m \Rightarrow \begin{cases} \dot{p}(t) = kp(t) + m, \\ p(0) = p_0, \end{cases} \quad t > 0,$$

Malthusian population model considering migration

$$\begin{cases} \dot{p}(t) = kp(t) + m, \\ p(0) = p_0, \end{cases} \quad t > 0, \quad k = b - d, \quad m = i - e.$$

There are two main approaches:

- *Unknown uncertainty*: Wiener process or Brownian motion. It requires the so-called Itô-calculus.
- *Known uncertainty*: It requires the so-called $L_p(\Omega)$ -calculus.

Itô-Stochastic Differential Equations (SDE's)

Assuming, for instance, that the birth-rate coefficient is affected by a Gaussian perturbation (*unknown uncertainty*):

$$\begin{aligned} \dot{p}(t) &= kp(t) + m, & t > 0, \\ p(0) &= p_0, \end{aligned} \quad \left. \right\}, \quad \textcolor{red}{k} \Rightarrow k + \lambda \underbrace{W'(t)}_{\text{white noise}}, \quad k \in \mathbb{R}, \lambda > 0,$$

$$\frac{dp(t)}{dt} = (k + \lambda W'(t))p(t) + m$$

$$dp(t) = (kp(t) + m)dt + \lambda p(t) \underbrace{W'(t)dt}_{dW(t)}$$

$$dp(t) = (kp(t) + m)dt + \lambda p(t)dW(t)$$

$$p(t) = p_0 + \int_0^t (kp(s) + m)dt + \underbrace{\int_0^t \lambda p(s)dW(s)}_{\text{Itô-type integral}} \xrightarrow{\text{Itô Lemma}} p(t)$$

Random Differential Equations (RDE's)

Known uncertainty:

- k is positive: $k \sim \text{Exp}(\lambda)$; $k \sim \text{Be}(\alpha; \beta)$.
- k is negative: $k \sim \text{Un}(-2, -0.5)$; $k \sim \text{N}(\mu; \sigma)$ truncated at $(-2, -0.5)$.

Malthusian population model considering migration

$$\begin{aligned}\dot{p}(t) &= kp(t) + m, & t > 0, \\ p(0) &= p_0,\end{aligned}\right\}, \quad k = b - d, \quad m = i - e.$$

In practice the birth, death, immigration, emigration rates and the initial population are fixed after sampling and measurements, hence it is more realistic to consider that:

k, m, p_0 are r.v.'s, defined in a common probability space, $(\Omega, \mathcal{F}, \mathbb{P})$
rather than deterministic constants



This motivates to consider the above model from a stochastic standpoint. As a consequence, its solution is a **stochastic process (s.p.)** rather than a classical function.



The main goals include to compute:

- The solution s.p.: $p(t) = p(t; \omega), \omega \in \Omega$.
- The mean function: $\mathbb{E}[p(t)]$.
- The variance function: $\mathbb{V}[p(t)]$.

To deal with RDE's, $L_p(\Omega)$ -calculus has demonstrate to be a powerful tool.

$p = 2 \Rightarrow$ mean square (m.s.) calculus

$$\begin{aligned} L_2(\Omega) &= \left\{ X : \Omega \rightarrow \mathbb{R}, \text{ 2-r.v.} \right\} \\ \|X\|_2 &= \left(\mathbb{E}[X^2] \right)^{1/2} < +\infty \end{aligned} \quad \left. \right\} \Rightarrow (L_2(\Omega), \|\cdot\|_2) \text{ Banach space}$$

- (Ω, \mathcal{F}, P) probability space
- $X : \Omega \rightarrow \mathbb{R}$ is a (continuous absolutely) real random variable (r.v.)
- F is a distribution function (d.f.); f is a probability density function (p.d.f.) of X
- X 2-r.v. $\Leftrightarrow \mathbb{E}[X^2] = \int_{\Omega} x^2 dF(\omega) = \int_{\mathbb{R}} x^2 f(x) dx < +\infty$
- X 2-r.v. $\Rightarrow \mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 < +\infty$
- Examples

mean square (m.s.) convergence of $\{X_n : n \geq 0\} \in L_2(\Omega)$

$$X_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} X \Leftrightarrow (\|X_n - X\|_2)^2 = \mathbb{E}[(X_n - X)^2] \xrightarrow{n \rightarrow \infty} 0$$

Some reasons to select mean square convergence

$$Z_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} Z \Rightarrow \begin{cases} \mathbb{E}[Z_n] \xrightarrow{n \rightarrow \infty} \mathbb{E}[Z], \\ \mathbb{V}[Z_n] \xrightarrow{n \rightarrow \infty} \mathbb{V}[Z]. \end{cases}$$



$$X_N(t) = \sum_{n=0}^N X_n t^n$$



$$t \in \mathcal{T} \text{ fixed}, Z_N = X_N(t) \Rightarrow \begin{cases} \mathbb{E}[X_N(t)] \xrightarrow{N \rightarrow \infty} \mathbb{E}[X(t)] \\ \mathbb{V}[X_N(t)] \xrightarrow{N \rightarrow \infty} \mathbb{V}[X(t)] \end{cases}$$

However, it would be more desirable to determine the **first probability density function** (1-p.d.f.), $f_1(p, t)$, associated to the solution s.p. $p(t)$ since from it one can compute, as merely particular cases, the mean and variance functions:

$$\mu_p(t) = \mathbb{E}[p(t)] = \int_{-\infty}^{\infty} p f_1(p, t) dp,$$

$$\sigma_p^2(t) = \mathbb{V}[p(t)] = \int_{-\infty}^{\infty} p^2 f_1(p, t) dp - (\mu_p(t))^2.$$

But in addition, from it one can also compute higher statistical moments:

$$\mathbb{E}[(p(t))^k] = \int_{-\infty}^{\infty} p^k f_1(p, t) dp, \quad k = 0, 1, 2, \dots,$$

and significant information such as the probability of the solution lies within a set of interest

$$\mathbb{P}[a \leq p(t) \leq b] = \int_a^b f_1(p, t) dp.$$

This improves the computation of rough bounds like

$$\mathbb{P}[|p(t) - \mu_p(t)| \geq \lambda] \leq \frac{(\sigma_p(t))^2}{\lambda^2},$$

usually used in practice.

The general random linear differential equation

Motivated by the previous presentation, in the following we focus on determining the 1-p.d.f., $f_Z(z, t)$, of the solution s.p. $Z(t)$ to the general linear random initial value problem (i.v.p.):

$$\begin{aligned}\dot{Z}(t) &= AZ(t) + B, \quad t > t_0, \\ Z(t_0) &= Z_0,\end{aligned}\right\}$$

where the data Z_0 , B and A are assumed to be absolutely continuous random variables (r.v.'s) defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose domains are assumed to be:

$$\begin{aligned}D_{Z_0} &= \{ z_0 = Z_0(\omega), \omega \in \Omega : z_{0,1} \leq z_0 \leq z_{0,2} \}, \\ D_B &= \{ b = B(\omega), \omega \in \Omega : b_1 \leq b \leq b_2 \}, \\ D_A &= \{ a = A(\omega), \omega \in \Omega : a_1 \leq a \leq a_2 \}.\end{aligned}$$

As we shall see later, the unifying element to conduct our study is the **Random Variable Transformation (R.V.T.) method**.

For the sake of clarity in the presentation we will distinguish the following cases:

TYPE	I.V.P.	CASE
H	$\begin{aligned}\dot{Z}(t) &= AZ(t) \\ Z(t_0) &= Z_0\end{aligned}\right\} \text{(I)}$	I.1 Z_0 is a random variable I.2 A is a random variable I.3 (Z_0, A) is a random vector
NH	$\begin{aligned}\dot{Z}(t) &= B \\ Z(t_0) &= Z_0\end{aligned}\right\} \text{(II)}$	II.1 Z_0 is a random variable II.2 B is a random variable II.3 (Z_0, B) is a random vector
	$\begin{aligned}\dot{Z}(t) &= AZ(t) + B \\ Z(t_0) &= Z_0\end{aligned}\right\} \text{(III)}$	III.1 Z_0 is a random variable III.2 B is a random variable III.3 A is a random variable III.4 (Z_0, B) is a random vector III.5 (Z_0, A) is a random vector III.6 (B, A) is a random vector III.7 (Z_0, B, A) is a random vector

$$Z(t) = e^{A(t-t_0)} Z_0 + \frac{B}{A} \left(e^{A(t-t_0)} - 1 \right), \quad t \geq t_0.$$

Remarks:

- I.V.P. (I): $\mathbb{P}[\{\omega \in \Omega : B(\omega) = 0\}] = 1$; I.V.P. (II): $\mathbb{P}[\{\omega \in \Omega : A(\omega) = 0\}] = 1$.
- Hereinafter, deterministic parameters will be written by lower case letters and r.v.'s by capital letters.
- Notation for the p.d.f.'s: $f_{Z_0}(z_0)$; $f_{Z_0,A}(z_0, a)$; $f_{Z_0,B,A}(z_0, b, a)$, etc.
- Standard and non-standard p.d.f.'s including copulas can be considered.

Preliminaries on Random Variable Transformation (R.V.T.) method

R.V.T. method: simple scalar version

H: Let X be a continuous r.v. with p.d.f. $f_X(x)$ with support $\mathcal{S}(X)$ and $Y = r(X)$ being r a bijective mapping.

T: The p.d.f. of Y , $g_Y(y)$, is given by:

$$g_Y(y) = f_X(x = s(y)) \left| \frac{ds(y)}{dy} \right|, \quad y \in \mathcal{S}(r(X)).$$

R.V.T. technique: general scalar version

H: Let X be a r.v. with p.d.f. $f_X(x)$ and codomain or support $D_X = \{x : f_X(x) > 0\}$. Let $Y = r(X)$ be a new r.v. generated by the map $r : \mathbb{R} \rightarrow \mathbb{R}$ which is assumed to be continuously differentiable on D_X and such that $r'(x) \neq 0$ except at a finite number of points. Let us suppose that for each $y \in \mathbb{R}$, there exist $m(y) \geq 1$ points: $x_1(y), x_2(y), \dots, x_{m(y)}(y) \in D_X$ such that

$$r(x_k(y)) = y, \quad r'(x_k(y)) \neq 0, \quad k = 1, 2, \dots, m(y).$$

T: Then
$$f_Y(y) = \begin{cases} \sum_{i=1}^{m(y)} f_X(x_k(y)) |r'(x_k(y))|^{-1} & \text{if } m(y) > 0, \\ 0 & \text{if } m(y) = 0. \end{cases}$$

Next, we shall present some particular cases of R.V.T. method that will be useful later.

Case I.1: $Z(t) = Z_0 e^{a(t-t_0)}$, $t \geq t_0$.

R.V.T. technique: linear transformation

- H**: Let X be a continuous r.v. with domain: $D_X = \{x : x_1 \leq x \leq x_2\}$ and p.d.f. $f_X(x)$.
- T**: Then, the p.d.f. $f_Y(y)$ of the linear transformation $Y = \alpha X + \beta$, $\alpha \neq 0$ is given by:

$$f_Y(y) = \frac{1}{|\alpha|} f_X\left(\frac{y-\beta}{\alpha}\right), \quad \text{where} \quad \begin{cases} y_1 = \alpha x_1 + \beta \leq y \leq \alpha x_2 + \beta = y_2 & \text{if } \alpha > 0, \\ y_1 = \alpha x_2 + \beta \leq y \leq \alpha x_1 + \beta = y_2 & \text{if } \alpha < 0. \end{cases}$$

If $\alpha = 0$, then $Y = \beta$ with probability 1 (w.p. 1) and

$$f_Y(y) = \delta(y - \beta), \quad -\infty < y < \infty,$$

where $\delta(\cdot)$ denotes the Dirac delta distribution.

Case I.2: $Z(t) = z_0 e^{\textcolor{blue}{A}(t-t_0)}$, $t \geq t_0$.

R.V.T. technique: exponential transformation

- H**: Let X be a continuous r.v. with domain: $D_X = \{x : x_1 \leq x \leq x_2\}$ and p.d.f. $f_X(x)$.
- T**: Then the p.d.f. $f_Y(y)$ of the exponential transformation $Y = \alpha e^{\beta X} + \gamma$, with $\alpha\beta \neq 0$ is given by:

$$f_Y(y) = \frac{1}{|\beta(y-\gamma)|} f_X\left(\frac{1}{\beta} \ln\left(\frac{y-\gamma}{\alpha}\right)\right),$$

where

$$\begin{cases} y_1 = \alpha e^{\beta x_1} + \gamma \leq y \leq \alpha e^{\beta x_2} + \gamma = y_2 & \text{if } \alpha\beta > 0, \\ y_1 = \alpha e^{\beta x_2} + \gamma \leq y \leq \alpha e^{\beta x_1} + \gamma = y_2 & \text{if } \alpha\beta < 0. \end{cases}$$

If $\alpha = 0$ or $\beta = 0$, then $Y = \alpha + \gamma$ w.p. 1 and

$$f_Y(y) = \delta(y - (\alpha + \gamma)), \quad -\infty < y < \infty.$$

Case I.3: $Z(t) = Z_0 e^{\mathcal{A}(t-t_0)}$, $t \geq t_0$.

R.V.T. technique: multi-dimensional version

H: Let $\mathbf{X} = (X_1, \dots, X_n)$ be a random vector of dimension n with joint p.d.f. $f_{\mathbf{X}}(\mathbf{x})$. Let $\mathbf{r}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a one-to-one deterministic map which is assumed to be continuous with respect to each one of its arguments, and with continuous partial derivatives.

T: Then, the joint p.d.f. $f_{\mathbf{Y}}(\mathbf{y})$ of the random vector $\mathbf{Y} = \mathbf{r}(\mathbf{X})$ is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{s}(\mathbf{y})) |J_n|,$$

where $\mathbf{s}(\mathbf{y})$ is the inverse transformation of $\mathbf{r}(\mathbf{x})$: $\mathbf{x} = \mathbf{r}^{-1}(\mathbf{y}) = \mathbf{s}(\mathbf{y})$ and J_n is the jacobian of the transformation, i.e.,

$$J_n = \det \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right) = \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial y_n} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix},$$

which is assumed to be different from zero.

Case II.3: $Z(t) = Z_0 + B(t - t_0)$, $t \geq t_0$.

R.V.T. technique: sum of two r.v.'s

H: Let (X_1, X_2) be a continuous random vector with joint p.d.f. $f_{X_1, X_2}(x_1, x_2)$ and respective domains: $D_{X_1} = \{x_1 : x_{1,1} \leq x_1 \leq x_{1,2}\}$ and $D_{X_2} = \{x_2 : x_{2,1} \leq x_2 \leq x_{2,2}\}$.

T: Then the p.d.f. $f_{Y_1}(y_1)$ of their sum $Y_1 = X_1 + X_2$ is given by:

$$f_{Y_1}(y_1) = \int_{x_{1,1}}^{x_{1,2}} f_{X_1, X_2}(x_1, y_1 - x_1) dx_1, \quad y_{1,1} = x_{1,1} + x_{2,1} \leq y_1 \leq x_{1,2} + x_{2,2} = y_{1,2},$$

or, equivalently by

$$f_{Y_1}(y_1) = \int_{x_{2,1}}^{x_{2,2}} f_{X_1, X_2}(y_1 - x_2, x_2) dx_2, \quad y_{1,1} = x_{1,1} + x_{2,1} \leq y_1 \leq x_{1,2} + x_{2,2} = y_{1,2}.$$

If X_1 and X_2 are independent r.v.'s, since $f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$, being $f_{X_i}(x_i)$ the p.d.f. of X_i , $i = 1, 2$, the p.d.f. of the sum of two independent r.v.'s is just the convolution of their respective p.d.f.'s:

$$f_{Y_1}(y_1) = \int_{x_{1,1}}^{x_{1,2}} f_{X_1}(x_1)f_{X_2}(y_1 - x_1) dx_1, \quad \text{or} \quad f_{Y_1}(y_1) = \int_{x_{2,1}}^{x_{2,2}} f_{X_1}(y_1 - x_2)f_{X_2}(x_2) dx_2.$$

Case I.3: $Z(t) = Z_0 e^{\textcolor{blue}{A}(t-t_0)}$, $t \geq t_0$.

R.V.T. technique: product of two r.v.'s

H: Let (X_1, X_2) be a continuous random vector with joint p.d.f. $f_{X_1, X_2}(x_1, x_2)$ with respective domains: $D_{X_1} = \{x_1 \neq 0 : x_{1,1} \leq x_1 \leq x_{1,2}\}$ and $D_{X_2} = \{x_2 : x_{2,1} \leq x_2 \leq x_{2,2}\}$.

T: Then the p.d.f. $f_{Y_1}(y_1)$ of their product $Y_1 = X_1 X_2$ is given by:

$$f_{Y_1}(y_1) = \int_{x_{1,1}}^{x_{1,2}} f_{X_1, X_2} \left(x_1, \frac{y_1}{x_1} \right) \frac{1}{|x_1|} dx_1.$$

Equivalently, if $D_{X_1} = \{x_1 : x_{1,1} \leq x_1 \leq x_{1,2}\}$ and $D_{X_2} = \{x_2 \neq 0 : x_{2,1} \leq x_2 \leq x_{2,2}\}$ then

$$f_{Y_1}(y_1) = \int_{x_{2,1}}^{x_{2,2}} f_{X_1, X_2} \left(\frac{y_1}{x_2}, x_2 \right) \frac{1}{|x_2|} dx_2. \quad (\star)$$

If X_1 and X_2 are independent r.v.'s with p.d.f.'s $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$, respectively, then previous formulas write:

$$f_{Y_1}(y_1) = \int_{x_{1,1}}^{x_{1,2}} f_{X_1}(x_1) f_{X_2} \left(\frac{y_1}{x_1} \right) \frac{1}{|x_1|} dx_1, \quad \text{or} \quad f_{Y_1}(y_1) = \int_{x_{2,1}}^{x_{2,2}} f_{X_1} \left(\frac{y_1}{x_2} \right) f_{X_2}(x_2) \frac{1}{|x_2|} dx_2,$$

respectively.

Computing the 1-p.d.f. of the solution s.p. of the general linear random differential equation: Some study-cases

Case I.1: Z_0 is a r.v.

In this case the solution s.p. has the following expression:

$$Z(t) = Z_0 e^{a(t-t_0)}, \quad t \geq t_0.$$

Next, we first fix $t: t \geq t_0$ and denote $Z = Z(t)$. Then, we apply R.V.T. method (linear transformation: $Y = \alpha X + \beta$, $\alpha \neq 0$) to:

$$\alpha = e^{a(t-t_0)} > 0, \quad \beta = 0, \quad X = Z_0, \quad Y = Z,$$

and, taking into account that $f_Y(y) = \frac{1}{|\alpha|} f_X\left(\frac{y-\beta}{\alpha}\right)$ and the domain of r.v. Z_0 , one gets:

$$f_1(z, t) = e^{-a(t-t_0)} f_{Z_0}\left(z e^{-a(t-t_0)}\right), \quad z_1 \leq z \leq z_2, \quad t \geq t_0,$$

where

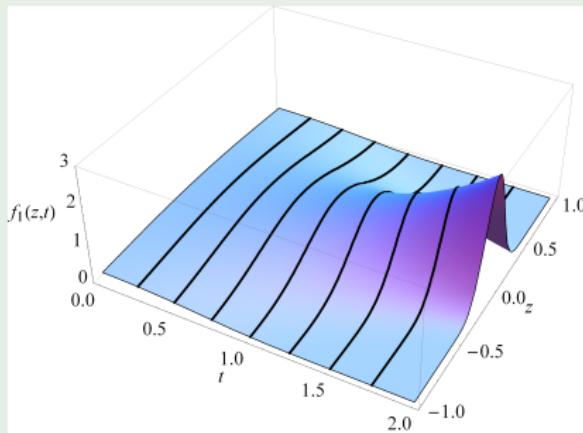
$$z_1 = z_{0,1} e^{a(t-t_0)}, \quad z_2 = z_{0,2} e^{a(t-t_0)}.$$

Example Case I.1: $Z_0 \sim N(\mu; \sigma^2)$, $\mu \in \mathbb{R}$ and $\sigma^2 > 0$

$$f_1(z, t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(a(t-t_0) + \frac{1}{2\sigma^2} (ze^{-a(t-t_0)} - \mu)^2\right)}, \quad -\infty < z < \infty, \quad t \geq t_0.$$

Example: $Z_0 \sim N(0; 1)$, $t_0 = 0$, $a = -1$.

$$Z(t) = Z_0 e^{-t}, \quad t \geq t_0.$$



Case I.2: A is a r.v.

In this case the solution s.p. has the following expression:

$$Z(t) = z_0 e^{A(t-t_0)}, \quad t \geq t_0.$$

Next, we first fix $t: t > t_0$ and denote $Z = Z(t)$. Then we apply R.V.T. method (exponential transformation: $Y = \alpha e^{\beta X} + \gamma$, $\alpha\beta \neq 0$) to:

$$\alpha = z_0 \neq 0, \quad \beta = t - t_0 \neq 0, \quad X = A, \quad \gamma = 0, \quad Y = Z.$$

Then, taking into account that $f_Y(y) = \frac{1}{|\beta(y-\gamma)|} f_X\left(\frac{1}{\beta} \ln\left(\frac{y-\gamma}{\alpha}\right)\right)$ and $z/z_0 = e^{\alpha(t-t_0)} > 0$ and the domain of r.v. A , one gets:

$$f_1(z, t) = \frac{1}{(t-t_0)|z|} f_A\left(\frac{1}{t-t_0} \ln\left(\frac{z}{z_0}\right)\right), \quad z_1 \leq z \leq z_2, \quad t > t_0,$$

where

$$\begin{aligned} z_1 &= z_0 e^{a_1(t-t_0)}, & z_2 &= z_0 e^{a_2(t-t_0)}, & \text{if } z_0 > 0, \\ z_1 &= z_0 e^{a_2(t-t_0)}, & z_2 &= z_0 e^{a_1(t-t_0)}, & \text{if } z_0 < 0. \end{aligned}$$

For $t = t_0$: $Z(t) = Z(t_0) = z_0$, which is deterministic. Then its 1-p.d.f. can be written by the Dirac delta function as follows:

$$f_1(z, t_0) = \delta(z - z_0), \quad -\infty < z < \infty.$$

Example Case I.2: $A \sim \text{Be}(\alpha; \beta)$, $\alpha, \beta > 0$ and $z_0 > 0$

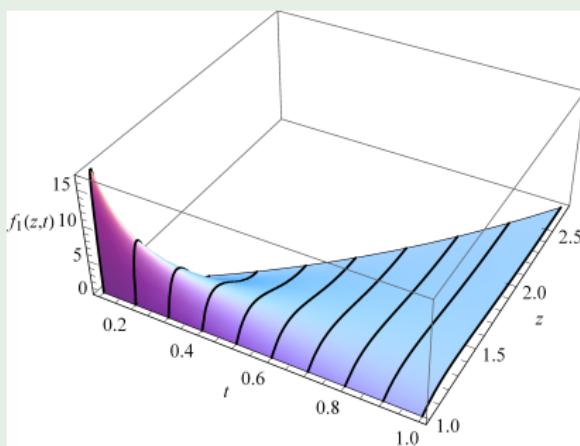
$$f_1(z, t) = \frac{1}{B(\alpha, \beta)|z|} \left(\frac{1}{t-t_0} \right)^\alpha \left(\ln \left(\frac{z}{z_0} \right) \right)^{\alpha-1} \left(1 - \frac{1}{t-t_0} \ln \left(\frac{z}{z_0} \right) \right)^{\beta-1},$$

$$z_0 \leq z \leq z_0 e^{t-t_0}, \quad t > t_0.$$

$$f_1(z, t_0) = \delta(z - z_0), \quad -\infty < z < \infty.$$

Remark: Since $z = z_0 e^{a(t-t_0)}$ and $0 \leq a \leq 1$, it is guaranteed that $0 \leq \frac{1}{t-t_0} \ln \left(\frac{z}{z_0} \right) \leq 1$.

Example: $A \sim \text{Be}(2; 3)$, $t_0 = 0$, $z_0 = 1$.



Case I.3: (Z_0, A) is a random vector

In this case that the solution s.p. has the following expression:

$$Z(t) = Z_1(t) Z_2(t), \quad \text{where} \quad \begin{cases} Z_1(t) &= Z_0, \\ Z_2(t) &= e^{A(t-t_0)}. \end{cases}$$

To compute the p.d.f. of $Z = Z(t)$, $t : t > t_0$ fix, first we will determine the joint p.d.f. of $Z_1 = Z_1(t)$ and $Z_2 = Z_2(t)$ by R.V.T. method ([two-dimensional version](#)) to:

$$\begin{array}{llllll} X_1 &= Z_0, & X_2 &= A, & r_1(z_0, a) &= z_0, & r_2(z_0, a) &= e^{a(t-t_0)}, \\ Y_1 &= Z_1, & Y_2 &= Z_2, & s_1(z_1, z_2) &= z_1, & s_2(z_1, z_2) &= \frac{\ln(z_2)}{t-t_0}. \end{array}$$

Hence, the Jacobian is given by

$$J_2 = \frac{\partial s_1(z_1, z_2)}{\partial z_1} \frac{\partial s_2(z_1, z_2)}{\partial z_2} = \frac{1}{z_2(t-t_0)} > 0,$$

therefore

$$f_{Z_1, Z_2}(z_1, z_2) = \frac{1}{z_2(t-t_0)} f_{Z_0, A} \left(z_1, \frac{\ln(z_2)}{t-t_0} \right), \quad z_{0,1} \leq z_1 \leq z_{0,2}, e^{a_1(t-t_0)} \leq z_2 \leq e^{a_2(t-t_0)}.$$

Now, we apply R.V.T. method ([product of two r.v.'s: \$Y_1 = X_1 X_2\$](#)) to obtain the p.d.f. of $Z = Z_1 Z_2$. As $Z_2 = e^{A(t-t_0)} \neq 0$, we will apply formula (\star) :

$$f_{Y_1}(y_1) = \int_{x_{2,1}}^{x_{2,2}} f_{X_1, X_2} \left(\frac{y_1}{x_2}, x_2 \right) \frac{1}{|x_2|} dx_2, \quad (\star)$$

to

$$X_1 = Z_1 = Z_0, \quad X_2 = Z_2 = e^{A(t-t_0)} > 0, \quad Y_1 = Z = Z_1 Z_2 :$$

$$\begin{aligned} f_1(z, t) = f_Z(z) &= \int_{z_{2,1}}^{z_{2,2}} f_{Z_1, Z_2} \left(\frac{z}{z_2}, z_2 \right) \frac{1}{z_2} dz_2 \\ &= \int_{z_{2,1}}^{z_{2,2}} f_{Z_0, A} \left(\frac{z}{z_2}, \frac{\ln(z_2)}{t-t_0} \right) \frac{1}{(z_2)^2(t-t_0)} dz_2, \quad z_1 \leq z \leq z_2, \quad t > t_0, \end{aligned}$$

where

$$z_{2,1} = e^{a_1(t-t_0)}, \quad z_{2,2} = e^{a_2(t-t_0)},$$

$$\begin{aligned} \hat{z}_1 &= z_{0,1} e^{a_1(t-t_0)}, & \hat{z}_2 &= z_{0,2} e^{a_2(t-t_0)}, & \text{if } z_{0,1} > 0, \\ \hat{z}_1 &= z_{0,1} e^{a_2(t-t_0)}, & \hat{z}_2 &= z_{0,2} e^{a_2(t-t_0)}, & \text{if } z_{0,1} z_{0,2} \leq 0, \\ \hat{z}_1 &= z_{0,1} e^{a_2(t-t_0)}, & \hat{z}_2 &= z_{0,2} e^{a_1(t-t_0)}, & \text{if } z_{0,2} < 0. \end{aligned}$$

$$f_1(z_0, t_0) = f_{Z_0}(z_0) = \int_{a_1}^{a_2} f_{Z_0, A}(z_0, a) da, \quad z_{0,1} \leq z_0 \leq z_{0,2}.$$

Example Case I.3: (Z_0, A) is a random vector whose components are independent

$$f_{Z_0, A}(z_0, a) = \begin{cases} 4az_0 & \text{if } 0 < z_0, a < 1, \\ 0 & \text{otherwise.} \end{cases}$$

we substitute into the obtained formula and after making some simplifications one obtains:

$$f_1(z, t) = \begin{cases} \frac{4z}{(t-t_0)^2} \int_1^{e^{t-t_0}} \frac{\ln(z_2)}{(z_2)^3} dz_2 & \text{if } 0 \leq z \leq 1, \\ \frac{4z}{(t-t_0)^2} \int_z^{e^{t-t_0}} \frac{\ln(z_2)}{(z_2)^3} dz_2 & \text{if } 1 \leq z \leq e^{t-t_0}, \end{cases} \quad t > t_0.$$

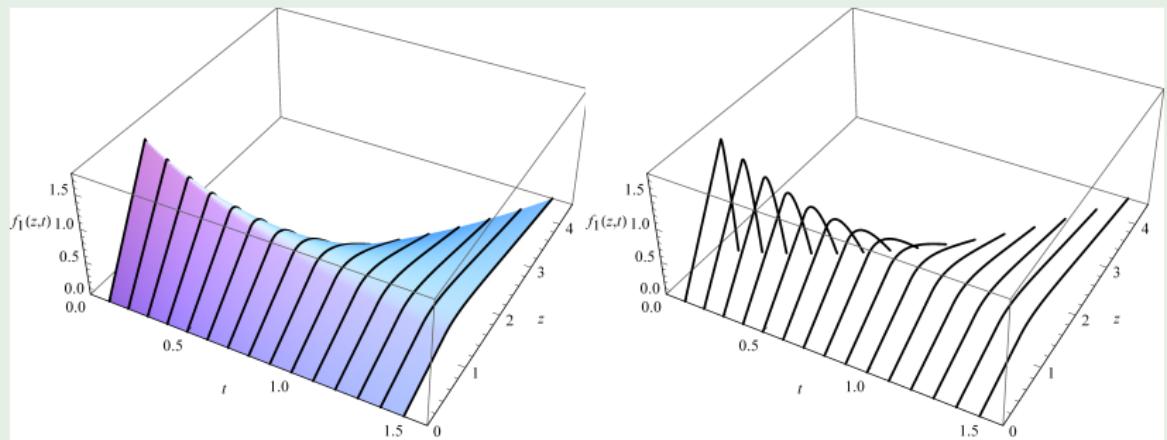
Let us take $t_0 = 0$. For $t > 0$:

$$f_1(z, t) = \begin{cases} \frac{z}{t^2} e^{-2t} (-1 + e^{2t} - 2t) & \text{if } 0 \leq z \leq 1, \\ \frac{z}{t^2} \left(-e^{-2t}(1+2t) + \frac{1+2\ln(z)}{z^2} \right) & \text{if } 1 \leq z \leq e^t. \end{cases} \quad t > 0.$$

For $t = 0$:

$$f_1(z_0, 0) = f_{Z_0}(z_0) = \int_0^1 4az_0 da = 2z_0, \quad z_{0,1} = 0 < z < 1 = z_{0,2}.$$

Example : $t > 0$, $t_0 = 0$.



Case II.3: (Z_0, B) is a random vector

In this case the solution s.p. has the following expression:

$$Z(t) = Z_1(t) + Z_2(t), \quad \text{where} \quad \begin{cases} Z_1(t) &= Z_0, \\ Z_2(t) &= B(t - t_0). \end{cases}$$

To compute the p.d.f. of $Z = Z(t)$, $t : t > t_0$ fix, first we will determine the joint p.d.f. of $Z_1 = Z_1(t)$ and $Z_2 = Z_2(t)$ by R.V.T. method ([two-dimensional version](#)) to:

$$\begin{array}{llll} X_1 = Z_0, & X_2 = B, & r_1(z_0, b) = z_0, & r_2(z_0, b) = b(t - t_0), \\ Y_1 = Z_1, & Y_2 = Z_2, & s_1(z_1, z_2) = z_1, & s_2(z_1, z_2) = \frac{z_2}{t - t_0}. \end{array}$$

Hence, the Jacobian is given by:

$$J_2 = \frac{\partial s_1(z_1, z_2)}{\partial z_1} \frac{\partial s_2(z_1, z_2)}{\partial z_2} = \frac{1}{t - t_0} > 0,$$

therefore

$$f_{Z_1, Z_2}(z_1, z_2) = \frac{1}{t - t_0} f_{Z_0, B}\left(z_1, \frac{z_2}{t - t_0}\right),$$

where

$$z_{1,1} = z_{0,1} \leq z_1 \leq z_{0,2} = z_{1,2}, \quad z_{2,1} = b_1(t - t_0) \leq z_2 \leq b_2(t - t_0) = z_{2,2}.$$

Now, we apply R.V.T. method ([sum of two r.v.'s: \$Y_1 = X_1 + X_2\$](#))

$$f_{Y_1}(y_1) = \int_{x_{1,1}}^{x_{1,2}} f_{X_1, X_2}(x_1, y_1 - x_1) dx_1,$$

to $X_1 = Z_1$, $X_2 = Z_2$ and $Y_1 = Z$ and we will obtain the p.d.f. of $Z = Z_1 + Z_2$:

$$\begin{aligned} f_1(z, t) = f_Z(z) &= \int_{z_{1,1}}^{z_{1,2}} f_{Z_1, Z_2}(z_1, z - z_1) dz_1, \\ &= \frac{1}{t - t_0} \int_{z_{0,1}}^{z_{0,2}} f_{Z_0, B}\left(z_0, \frac{z - z_0}{t - t_0}\right) dz_0, \quad \hat{z}_1 \leq z \leq \hat{z}_2, \quad t > t_0, \end{aligned}$$

where

$$\hat{z}_1 = z_{0,1} + b_1(t - t_0) \leq z \leq z_{0,2} + b_2(t - t_0) = \hat{z}_2.$$

If $t = t_0$, then $Z(t) = Z(t_0) = Z_0$ and the 1-p.d.f. is the Z_0 -marginal p.d.f.:

$$f_1(z_0, t_0) = f_{Z_0}(z_0) = \int_{b_1}^{b_2} f_{Z_0, B}(z_0, b) db, \quad z_{0,1} \leq z_0 \leq z_{0,2}.$$

Example Case II.3: (Z_0, B) is a random vector whose components are dependent

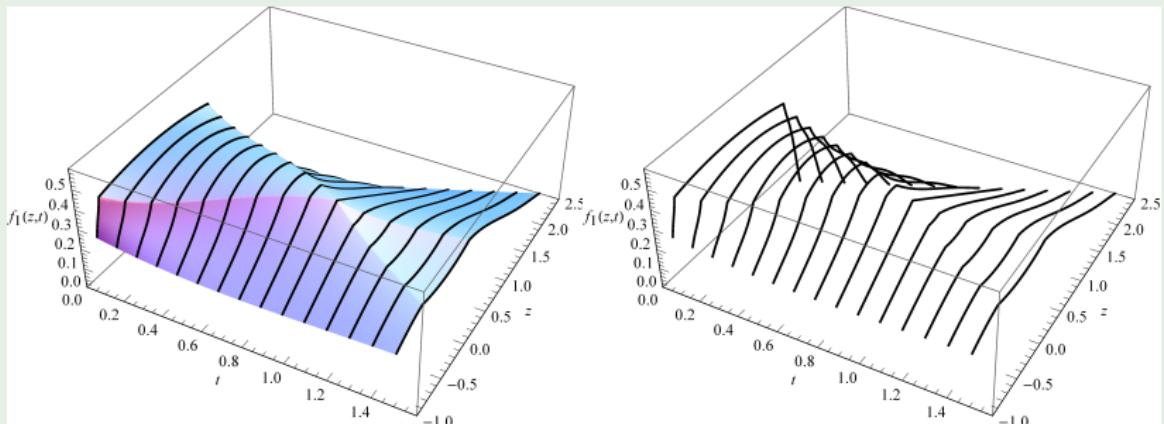
$$f_{Z_0, B}(z_0, b) = \begin{cases} \frac{1}{4} + \frac{1}{4}(z_0)^3 b - \frac{1}{4} z_0 b^3 & \text{if } -1 < z_0 < 1, -1 < b < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example : $t > 0$, $t_0 = 0$.

We substitute into the obtained formula and after making some simplifications to obtain:

$$f_1(z, t) = \frac{1}{t} \int_{\max\{z-t, -1\}}^{\min\{1, z+t\}} \left(\frac{1}{4} + \frac{1}{4}(z_0)^3 \left(\frac{z-z_0}{t} \right) - \frac{1}{4} z_0 \left(\frac{z-z_0}{t} \right)^3 \right) dz_0,$$

$$-1-t \leq z \leq 1+t.$$



Case III.1: Z_0 is a r.v. Here, we illustrate the computation of the mean, variance and probabilities of interest

In this case the solution s.p. has the following expression:

$$Z(t) = e^{a(t-t_0)} Z_0 + \frac{b}{a} \left(e^{a(t-t_0)} - 1 \right), \quad t \geq t_0.$$

Next, we first fix $t : t \geq t_0$ and denote $Z = Z(t)$. Then we apply R.V.T. method (linear transformation: $Y = \alpha X + \beta$, $\alpha \neq 0$) to:

$$\alpha = e^{a(t-t_0)} > 0, \quad \beta = \frac{b}{a} \left(e^{a(t-t_0)} - 1 \right), \quad X = Z_0, \quad Y = Z.$$

and, taking into account that $f_Y(y) = \frac{1}{|\alpha|} f_X\left(\frac{y-\beta}{\alpha}\right)$ the domain of r.v. Z_0 , one gets:

$$f_1(z, t) = e^{-a(t-t_0)} f_{Z_0} \left(e^{-a(t-t_0)} \left(z + \frac{b}{a} \right) - \frac{b}{a} \right), \quad z_1 \leq z \leq z_2, \quad t \geq t_0,$$

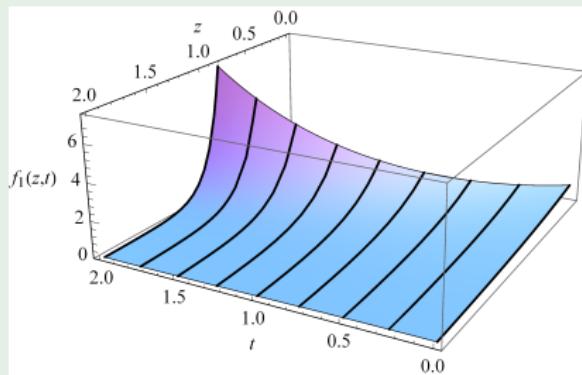
where

$$z_1 = z_{0,1} e^{a(t-t_0)} + \frac{b}{a} \left(e^{a(t-t_0)} - 1 \right), \quad z_2 = z_{0,2} e^{a(t-t_0)} + \frac{b}{a} \left(e^{a(t-t_0)} - 1 \right).$$

Example Case III.1: $Z_0 \sim \text{Exp}(\lambda)$, $\lambda > 0$

$$f_1(z, t) = \lambda e^{-\left(a(t-t_0)+\lambda\left((z+\frac{b}{a})e^{-a(t-t_0)}-\frac{b}{a}\right)\right)}, \quad \frac{b}{a} \left(e^{a(t-t_0)} - 1\right) \leq z < +\infty, \quad t \geq t_0,$$

Example : $\lambda = 1$, $t_0 = 0$ $a = -1$, $b = 1$.



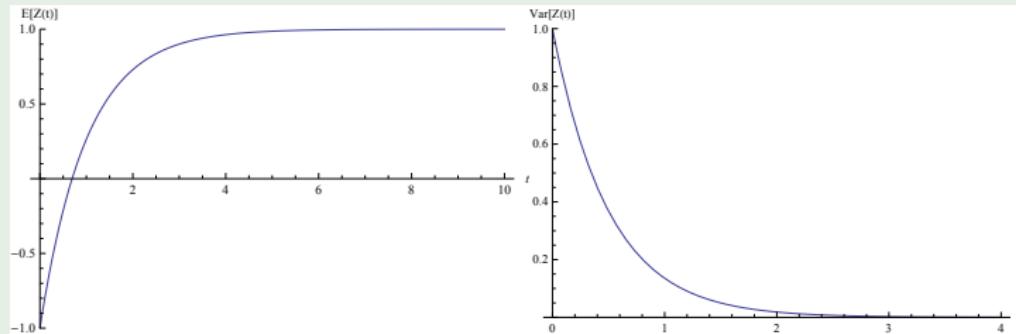
Example Case III.1: Computing some statistical properties by the 1-p.d.f.

Moments w.r.t. the origin:

$$\alpha_n(t) = \mathbb{E}[(Z(t))^k] = \int_{\frac{b}{a}(e^{at}-1)}^{\infty} z^k f_1(z, t) dz, \quad k = 0, 1, 2, \dots$$

$$\begin{aligned}\mathbb{E}[Z(t)] &= \alpha_1(t) &= \frac{-b\lambda + e^{at}(a + b\lambda)}{\lambda a}, \\ \mathbb{V}[Z(t)] &= \alpha_2(t) - (\alpha_1(t))^2 &= \frac{e^{2at}}{\lambda^2}.\end{aligned}$$

Example: $\lambda = 1$, $t_0 = 0$, $a = -1$, $b = 1$.



The computation of probabilities can also be carried out directly through the 1-p.d.f. For instance, it may be of interest to determine the probability that the solution lies between two fixed values, say, $v_1 = 2$ and $v_2 = 3$:

$$\begin{aligned}\mathbb{P}[2 \leq Z \leq 3] &= \int_2^3 f_1(z, t) dz \\ &= -e^{\frac{\lambda}{a} \left(b - (3a + b)e^{a(-t+t_0)} \right)} + e^{\frac{\lambda}{a} \left(b - e^{a(-t+t_0)} \left(b + a \text{Max} \left[2, \frac{b(-1+e^{a(t-t_0)})}{a} \right] \right) \right)}.\end{aligned}$$

Some important remarks regarding the application of R.V.T. technique: limitations and possibilities

Remark 1: The importance of making an appropriate choice

Let us consider Case III.5. If we write the solution s.p. in the following form:

$$Z(t) = Z_1(t) + Z_2(t), \quad \text{where} \quad \begin{cases} Z_1(t) &= Z_0 e^{A(t-t_0)}, \\ Z_2(t) &= \frac{b}{A} \left(e^{A(t-t_0)} - 1 \right), \end{cases}$$

then the application of R.V.T. (two-dimensional version) with the following choice:

$$\begin{array}{llll} X_1 = Z_0, & X_2 = A, & r_1(z_0, a) = z_0 e^{a(t-t_0)}, & r_2(a) = \frac{b}{a} \left(e^{a(t-t_0)} - 1 \right), \\ Y_1 = Z_1, & Y_2 = Z_2, & s_1(z_1, z_2) = ? & s_2(z_2) = ? \end{array}$$

does not lead to fruitful results since we cannot isolate $z_0 = s_1(z_1, z_2)$ and $a = s_2(z_1, z_2)$ and this would ruin our goal.

Notice that the previous drawback can be overcome as follows:

$$Z(t) = Z_1(t) + Z_2(t), \quad \text{where} \quad \begin{cases} Z_1(t) &= \left(Z_0 + \frac{b}{A}\right) e^{A(t-t_0)}, \\ Z_2(t) &= -\frac{b}{A}. \end{cases}$$

and applying R.V.T. (two-dimensional version) with the following choice:

$$\begin{array}{lll} X_1 = Z_0, & X_2 = A, & r_1(z_0, a) = \left(z_0 + \frac{b}{a}\right) e^{a(t-t_0)}, \\ Y_1 = Z_1, & Y_2 = Z_2, & s_1(z_1, z_2) = z_1 e^{\frac{b}{z_2}(t-t_0)} + z_2, \\ & & s_2(z_2) = -\frac{b}{z_2}. \end{array}$$

However, sometimes a good choice is not enough to apply R.V.T. method. Take a meanwhile to deal with the *apparent* simplest Case III.3 where:

$$Z = r(A), \quad \text{where } r(A) = z_0 e^{A(t-t_0)} + \frac{b}{A} \left(e^{A(t-t_0)} - 1 \right).$$

Can you isolate the r.v. A ?

H: Suppose z is defined as a function of the variable a by an equation of the form: $z = r(a)$ where r is analytic about the point a_0 where $r'(a_0) \neq 0$.

T: Then, it is possible to invert (or to solve) the equation for a : $a = s(z)$ on a neighbourhood $\mathcal{N}(r(a_0); \delta)$, $\delta > 0$ of $r(a_0)$:

$$a = s(z) = a_0 + \sum_{n=1}^{\infty} \left(\lim_{a \rightarrow a_0} \left(\frac{d^{n-1}}{da^{n-1}} \left(\frac{a - a_0}{r(a) - r(a_0)} \right)^n \right) \frac{(z - r(a_0))^n}{n!} \right), \quad z \in \mathcal{N}(r(a_0); \delta), \delta > 0.$$

- **Step 1:** Divide the domain of the map r (or equivalently, the domain of the r.v. A) into k subintervals: $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_k$ where r is monotone.
- **Step 2:** For every subinterval \mathcal{A}_j , $1 \leq j \leq k$, select $a_{0,j} \in \mathcal{A}_j$ such that $r'(a_{0,j}) \neq 0$. By Lagrange–Bürmann formula, construct the inverse, say $s_j(z)$, of the map $r(a) = r_j(a)$ on \mathcal{A}_j :

$$s_j(z) = a_{0,j} + \sum_{n=1}^{\infty} \left(\lim_{a \rightarrow a_{0,j}} \left(\frac{d^{n-1}}{da^{n-1}} \left(\frac{a - a_{0,j}}{r(a) - r(a_{0,j})} \right)^n \right) \frac{(z - r(a_{0,j}))^n}{n!} \right), \quad z \in \mathcal{N}(r(a_{0,j}); \delta).$$

- **Step 3:** Compute the derivative of $s_j(z)$:

$$\frac{ds_j(z)}{dz} = \sum_{n=1}^{\infty} \left(\lim_{a \rightarrow a_{0,j}} \left(\frac{d^{n-1}}{da^{n-1}} \left(\frac{a - a_{0,j}}{r(a) - r(a_{0,j})} \right)^n \right) \frac{(z - r(a_{0,j}))^{n-1}}{(n-1)!} \right), \quad z \in \mathcal{N}(r(a_{0,j}); \delta).$$

- **Step 4:** Construct the 1-p.d.f. of $Z(t)$ as follows:

$$f_1(z, t) = \sum_{j=1}^k f_A(s_j(z)) \left| \frac{ds_j(z)}{dz} \right|.$$

Often, the above infinite series must be truncated at the term N_j to control computational burden:

$$s_{j,N_j}(z) = a_{0,j} + \sum_{n=1}^{N_j} \left(\lim_{a \rightarrow a_{0,j}} \left(\frac{d^{n-1}}{da^{n-1}} \left(\frac{a - a_{0,j}}{r(a) - r(a_{0,j})} \right)^n \right) \frac{(z - r(a_{0,j}))^n}{n!} \right).$$

Thus, an approximation of its derivative is:

$$\frac{ds_{j,N_j}(z)}{dz} = \sum_{n=1}^{N_j} \left(\lim_{a \rightarrow a_{0,j}} \left(\frac{d^{n-1}}{da^{n-1}} \left(\frac{a - a_{0,j}}{r(a) - r(a_{0,j})} \right)^n \right) \frac{(z - r(a_{0,j}))^{n-1}}{(n-1)!} \right).$$

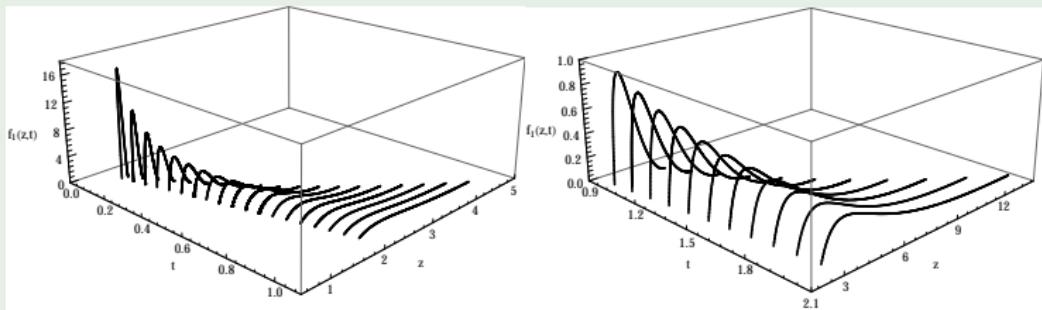
Repeating the foregoing process on each interval \mathcal{A}_j , $1 \leq j \leq k$, one gets the corresponding approximation of $f_1(z, t)$ given by:

$$f_1(z, t) = \sum_{j=1}^k f_A(s_{j,N_j}(z)) \left| \frac{ds_{j,N_j}(z)}{dz} \right|.$$

Example Case III.3: $A \sim \text{Be}(\alpha = 2; \beta = 3)$, $b = 1$, $t_0 = 0$, $z_0 = 1$

Since in this case $r(A)$ is monotone on the whole interval $\mathcal{A}_1 = [0, 1]$, we take $k = 1$. In order to carry out computations, \mathcal{A}_1 has been split into 7 subintervals in accordance with the process described previously. In each subinterval, an approximation of degree $N_j = 2$, $1 \leq j \leq 7$, has been used.

For the sake of clarity in the representation, due to differences in the scale the plot has been split in two pieces: $t \in [0, 1]$ and $t \in [1, 2]$.



Remark 2: Computing the 2-p.d.f. of the solution s.p.

Let us consider Case II.3 and let us fix t_1, t_2 such as $t_2 > t_1 \geq t_0$ and denote $Z_1 = Z(t_1)$ and $Z_2 = Z(t_2)$. All we need to determine the 2-p.d.f. of the solution s.p. $Z(t)$ is computing the joint p.d.f. of r.v.'s Z_1 and Z_2 . Notice that:

$$Z(t) = Z_0 + B(t - t_0).$$

Then, we apply R.V.T. (two-dimensional version) with the following choice:

$$\begin{aligned} X_1 &= Z_0, & X_2 &= B, & r_1(z_0, b) &= z_0 + b(t_1 - t_0), & r_2(z_0, b) &= z_0 + b(t_2 - t_0), \\ Y_1 &= Z_1, & Y_2 &= Z_2, & s_1(z_1, z_2) &= \frac{z_1(t_2 - t_0) - z_2(t_1 - t_0)}{t_2 - t_1}, & s_2(z_1, z_2) &= \frac{z_2 - z_1}{t_2 - t_1}, \end{aligned}$$

Now, taking into account that:

$$\frac{ds_1(z_1, z_2)}{dz_1} = \frac{t_2 - t_0}{t_2 - t_1}, \quad \frac{ds_1(z_1, z_2)}{dz_2} = -\frac{t_1 - t_0}{t_2 - t_1},$$

$$\frac{ds_2(z_1, z_2)}{dz_1} = -\frac{1}{t_2 - t_1}, \quad \frac{ds_2(z_1, z_2)}{dz_2} = \frac{1}{t_2 - t_1},$$

one obtains the Jacobian:

$$|J_2| = \frac{1}{t_2 - t_1} > 0.$$

Finally:

$$\begin{aligned}f_2(z_1, t_1; z_2, t_2) &= f_{Z_1, Z_2}(z_1, z_2) \\&= f_{Z_0, B}\left(\frac{z_1(t_2 - t_0) - z_2(t_1 - t_0)}{t_2 - t_1}, \frac{z_2 - z_1}{t_2 - t_1}\right) \frac{1}{t_2 - t_1},\end{aligned}$$

where $z_{1,1} \leq z_1 \leq z_{1,2}$, $z_{2,1} \leq z_2 \leq z_{2,2}$ satisfy

$$z_{1,1} = z_{0,1} + b_1(t_1 - t_0), \quad z_{1,2} = z_{0,2} + b_2(t_1 - t_0),$$

$$z_{2,1} = z_{0,1} + b_1(t_2 - t_0), \quad z_{2,2} = z_{0,2} + b_2(t_2 - t_0).$$

From the 2-p.d.f., we can calculate relevant probabilistic properties such as the correlation function:

$$\Gamma_Z(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 z_2 f_2(z_1, t_1; z_2, t_2) dz_1 dz_2.$$

Remark 3: Sometimes the computation of the 1-p.d.f. gives full information of the solution s.p.

Let us consider Case III.1 for which:

$$Z(t) = \left(Z_0 + \frac{b}{a} \right) e^{a(t-t_0)} - \frac{b}{a}.$$

In this case, the solution s.p. at t_2 can be represented as follows:

$$\begin{aligned} Z(t_2) &= \left(Z_0 + \frac{b}{a} \right) e^{a(t_2-t_0)} - \frac{b}{a}, \\ &= e^{a(t_2-t_1)} \left(Z_0 + \frac{b}{a} \right) e^{a(t_1-t_0)} - \frac{b}{a} \\ &= e^{a(t_2-t_1)} \left(Z(t_1) + \frac{b}{a} \right) - \frac{b}{a} \\ &= e^{a(t_2-t_1)} Z(t_1) + \frac{b}{a} \left(e^{a(t_2-t_1)} - 1 \right). \end{aligned}$$

From this expression we see that the behaviour of the solution $Z(t)$ at the time instant t_2 is deterministically given by a linear transformation of $Z(t_1)$. Therefore, *the computation of the 2-p.d.f. is not required*.

Let us check it from another point of view!

Let us assume without loss of generality that the expectation of the initial condition is zero: $\mathbb{E}[Z_0] = 0$ and its variance is $\sigma_{Z_0}^2 > 0$. Then it is easy to check that:

$$\begin{aligned}\mathbb{E}[Z(t_i)] &= \frac{b}{a} e^{a(t_i - t_0)} - \frac{b}{a}, \quad i = 1, 2, \\ \sigma_{Z(t_i)}^2 &= \sigma_{Z_0}^2 e^{2a(t_i - t_0)}, \quad i = 1, 2, \\ \mathbb{E}[Z(t_1)Z(t_2)] &= \sigma_{Z_0}^2 e^{a(t_2 + t_1 - 2t_0)} + \left(\frac{b}{a}\right)^2 \left(e^{a(t_2 + t_1 - 2t_0)} - e^{a(t_2 - t_0)} - e^{a(t_1 - t_0)} + 1\right).\end{aligned}$$

Then the correlation coefficient function is given by

$$\rho_{Z(t_1), Z(t_2)} = \frac{\mathbb{E}[Z(t_1)Z(t_2)] - \mathbb{E}[Z(t_1)]\mathbb{E}[Z(t_2)]}{\sigma_{Z(t_1)}\sigma_{Z(t_2)}} = 1.$$

$Z(t_2)$ is completely determined by $Z(t_1)$!

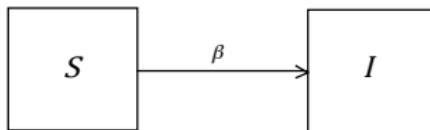
Remark 4: Different but equivalent representations of the 1-p.d.f.

It is important to underline that there usually exist several ways to conduct the study when applying R.V.T. method, although some of them are easier. Therefore, different *apparently* results can appear.

Part III

Nonlinear Models in Epidemiology

Motivating the nonlinear case: The SI-type epidemiological model



- $S(t)$ number of susceptibles in the time instant t .
- $I(t)$ number of infected in the time instant t .
- n size of the total population. It is assumed to be constant for all time t .
- $\beta > 0$ rate of decline in the number of susceptibles.

$$\begin{cases} S'(t) &= -\frac{\beta}{n} S(t)[n - S(t)], \quad t > 0, \\ S(0) &= m, \end{cases}$$

Putting the change of variable: $P(t) = \frac{S(t)}{n} \in [0, 1]$, the model can be recast as follows

$$\begin{cases} P'(t) &= -\beta P(t)[1 - P(t)], \quad t > 0, \\ P(0) &= P_0 = m/n. \end{cases}$$

normalized SI-type epidemiological model

$$\begin{cases} P'(t) &= -\beta P(t)[1 - P(t)], \quad t > 0, \\ P(0) &= P_0. \end{cases}$$

It is more realistic to assume that β and P_0 are r.v.'s rather than deterministic constants. We will assume that they are independent r.v.'s with p.d.f.'s $f_{P_0}(p_0)$ and $f_\beta(\beta)$ and domains

$$\begin{aligned} D_{P_0} &= \{ p_0 = P_0(\omega), \omega \in \Omega : 0 \leq p_{0,1} \leq p_0 \leq p_{0,2} \leq 1 \}, \\ D_\beta &= \{ \beta = \beta(\omega), \omega \in \Omega : 0 < \beta_1 < \beta < \beta_2 \}, \end{aligned}$$

respectively.

To compute the 1-p.d.f. of the solution s.p. $P(t)$. To this end, we make several changes of variables to accommodate the nonlinear SI-model to random linear model previously studied using the linearization technique:

- First change of variable: $Q(t) = \frac{1}{P(t)}$. Then, the problem SI-model writes

$$\begin{aligned} Q'(t) &= \beta Q(t) - \beta, & t > 0, \\ Q(0) &= \frac{1}{P_0}. \end{aligned} \quad \left. \right\}$$

- Second change of variable: $H(t) = Q(t) - 1$. This leads

$$\begin{aligned} H'(t) &= \beta H(t), & t > 0, \\ H(0) &= \frac{1}{P_0} - 1. \end{aligned} \quad \left. \right\}$$

Using R.V.T. technique one can establish the following result:

Case I.3 of linear random model

H: Let us consider the linear random i.v.p.

$$\begin{aligned}\dot{Z}(t) &= AZ(t), \quad t > t_0, \\ Z(t_0) &= Z_0,\end{aligned}\quad \left. \begin{array}{l} Z_0, A \text{ r.v.'s with joint p.d.f. } f_{Z_0,A}(z_0, a) \end{array} \right\} \quad (1)$$

and domains

$$D_{Z_0} = \{z_0 = Z_0(\omega), \omega \in \Omega : z_{0,1} \leq z_0 \leq z_{0,2}\}, \quad D_A = \{a = A(\omega), \omega \in \Omega : a_1 \leq a \leq a_2\}.$$

T: Then, the 1-p.d.f. of the solution s.p. $Z(t)$ of (1) is given by

$$f_1(z, t) = \frac{1}{t - t_0} \int_{e^{a_1(t-t_0)}}^{e^{a_2(t-t_0)}} f_{Z_0,A} \left(\frac{z}{\xi}, \frac{\ln(\xi)}{t - t_0} \right) \frac{1}{\xi^2} d\xi, \quad z_1 \leq z \leq z_2, \forall t > t_0,$$

where

$$\begin{aligned}z_1 &= z_{0,1} e^{a_1(t-t_0)}, & z_2 &= z_{0,2} e^{a_2(t-t_0)}, & \text{if } z_{0,1} > 0, \\z_1 &= z_{0,1} e^{a_2(t-t_0)}, & z_2 &= z_{0,2} e^{a_2(t-t_0)}, & \text{if } z_{0,1} z_{0,2} \leq 0, \\z_1 &= z_{0,1} e^{a_2(t-t_0)}, & z_2 &= z_{0,2} e^{a_1(t-t_0)}, & \text{if } z_{0,2} < 0.\end{aligned}$$

If $t = t_0$,

$$f_1(z_0, t_0) = \int_{a_1}^{a_2} f_{Z_0,A}(z_0, a) da, \quad z_{0,1} \leq z_0 \leq z_{0,2}.$$

We identify the inputs of both problems:

$$\left. \begin{array}{rcl} H'(t) & = & \beta H(t), \quad t > 0, \\ H(0) & = & \frac{1}{P_0} - 1. \end{array} \right\} \equiv \left. \begin{array}{rcl} \dot{Z}(t) & = & AZ(t), \quad t > t_0, \\ Z(t_0) & = & Z_0, \end{array} \right\}$$

$$Z_0 = \frac{1}{P_0} - 1, \quad A = \beta, \quad Z(t) = H(t), \quad t_0 = 0,$$

and, fixed $t > 0$, the p.d.f. of r.v. $H = H(t)$ yields

$$f_H(h) = \frac{1}{t} \int_{e^{a_1 t}}^{e^{a_2 t}} f_{Z_0, A} \left(\frac{h}{\xi}, \frac{\ln(\xi)}{t} \right) \frac{1}{\xi^2} d\xi = \frac{1}{t} \int_{e^{a_1 t}}^{e^{a_2 t}} f_{Z_0} \left(\frac{h}{\xi} \right) f_A \left(\frac{\ln(\xi)}{t} \right) \frac{1}{\xi^2} d\xi,$$

where independence between r.v.'s Z_0 and A has been used.

Now, we need to write $f_H(h)$ in terms of the p.d.f.'s of the inputs P_0 and β . With this aim we establish the following specialization of R.V.T. method:

H: Let $c \in \mathbb{R}$ and X be an absolutely continuous real r.v. defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, with p.d.f. $f_X(x)$. Assume that X is a non-zero r.v. and let us denote by D_X the domain of r.v. X , where

$$D_X = I_x^- \cup I_x^+, \quad \begin{cases} I_x^- = \{x = X(\omega) \in \mathbb{R} : -\infty < x < 0, \omega \in \Omega\}, \\ I_x^+ = \{x = X(\omega) \in \mathbb{R} : 0 < x < +\infty, \omega \in \Omega\}. \end{cases}$$

T: Then, the p.d.f. $f_Y(y)$ of the inverse-vertical translation transformation $Y = \frac{1}{X} + c$ is given by

$$f_Y(y) = \frac{1}{(y-c)^2} f_X\left(\frac{1}{y-c}\right), \quad y \in D_Y = I_y^- \cup I_y^+, \quad \begin{cases} I_y^- = \{y \in \mathbb{R} : y < c\}, \\ I_y^+ = \{y \in \mathbb{R} : y > c\}. \end{cases}$$

Applying this result to:

$$X = P_0, Y = Z_0, c = -1, \quad \left(Z_0 = \frac{1}{P_0} - 1\right)$$

one gets

$$f_H(h) = \frac{1}{t} \int_{e^{a_1 t}}^{e^{a_2 t}} f_{Z_0}\left(\frac{h}{\xi}\right) f_A\left(\frac{\ln(\xi)}{t}\right) \frac{1}{\xi^2} d\xi = \frac{1}{t} \int_{e^{\beta_1 t}}^{e^{\beta_2 t}} f_{P_0}\left(\frac{\xi}{h+\xi}\right) f_\beta\left(\frac{\ln(\xi)}{t}\right) \frac{1}{(h+\xi)^2} d\xi.$$

Remember that: $P(t) = \frac{1}{H(t)+1}$, so fixed t , we finally need to recover the p.d.f. $f_P(p)$ of r.v. $P = P(t)$ from the p.d.f. $f_H(h)$. With this end, we establish the following result:

R.V.T. technique: inverse-horizontal translation transformation

H: Let $d \in \mathbb{R}$ and X be an absolutely continuous real r.v. defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, with p.d.f. $f_X(x)$. Assume that $X - d$ is a non-zero r.v. and let us denote by D_X the domain of r.v. X , where

$$D_X = I_x^- \cup I_x^+, \quad \begin{cases} I_x^- = \{x = X(\omega) \in \mathbb{R} : -\infty < x < d, \omega \in \Omega\}, \\ I_x^+ = \{x = X(\omega) \in \mathbb{R} : d < x < +\infty, \omega \in \Omega\}. \end{cases}$$

T: Then, the p.d.f. $f_Y(y)$ of the inverse-horizontal translation transformation $Y = \frac{1}{X-d}$ is given by

$$f_Y(y) = \frac{1}{y^2} f_X \left(\frac{1}{y} + d \right), \quad y \in D_Y = I_y^- \cup I_y^+, \quad \begin{cases} I_y^- = \{y \in \mathbb{R} : y < 0\}, \\ I_y^+ = \{y \in \mathbb{R} : y > 0\}. \end{cases}$$

Applying this result to:

$$X = H, Y = P, d = -1,$$

one gets

$$f_P(p) = \frac{1}{p^2} f_H \left(\frac{1}{p} - 1 \right) = \frac{1}{t} \int_{e^{\beta_1 t}}^{e^{\beta_2 t}} f_{P_0} \left(\frac{p\xi}{1-p+p\xi} \right) f_\beta \left(\frac{\ln(\xi)}{t} \right) \frac{1}{(1-p+p\xi)^2} d\xi.$$

Summarizing,

1-p.d.f. of the solution s.p. of the normalized SI-type epidemiological model

H: Let us consider the random i.v.p.:

$$\begin{cases} P'(t) &= -\beta P(t)[1 - P(t)], \quad t > 0, \\ P(0) &= P_0, \end{cases}$$

where β and P_0 are independent r.v.'s with p.d.f.'s $f_{P_0}(p_0)$ and $f_\beta(\beta)$ and domains

$$\begin{aligned} D_{P_0} &= \{p_0 = P_0(\omega), \omega \in \Omega : 0 \leq p_{0,1} \leq p_0 \leq p_{0,2} \leq 1\}, \\ D_\beta &= \{\beta = \beta(\omega), \omega \in \Omega : 0 \leq \beta_1 < \beta < \beta_2\}, \end{aligned}$$

respectively

T: Then, the 1-p.d.f. of the solution s.p. $P(t)$ is given by:

$$f_1(p, t) = \begin{cases} \frac{1}{t} \int_{e^{\beta_1 t}}^{e^{\beta_2 t}} f_{P_0} \left(\frac{p\xi}{1 - p + p\xi} \right) f_\beta \left(\frac{\ln(\xi)}{t} \right) \frac{1}{(1 - p + p\xi)^2} d\xi & \text{if } t > 0, \\ f_{P_0}(p_0) & \text{if } t = 0. \end{cases}$$

From this 1-p.d.f. important information related to SI-epidemiological model can be computed straightforwardly:

- **Mean and variance:**

$$\mu_P(t) = \mathbb{E}[P(t)] = \int_{-\infty}^{\infty} pf_1(p, t) dp, \quad (\sigma_P(t))^2 = \mathbb{V}[P(t)] = \int_{-\infty}^{\infty} p^2 f_1(p, t) dp - (\mu_P(t))^2,$$

- **Bounds for probabilities upon intervals of interest and more:**

$$\mathbb{P}[|P(t) - \mu_P(t)| \geq \lambda] \leq \frac{(\sigma_P(t))^2}{\lambda^2},$$

$$\mathbb{P}[a \leq P(t) \leq b] = \int_a^b f_1(p, t) dp,$$

- **Confidence intervals:** Fixed $\alpha \in (0, 1)$, for each time instant t one can determine $x_1(t)$ and $x_2(t)$, such that

$$1 - \alpha = \mathbb{P}(\{\omega \in \Omega : P(t; \omega) \in [x_1(t), x_2(t)]\}) = \int_{x_1(t)}^{x_2(t)} f_1(p, t) dp,$$

and

$$\int_0^{x_1(t)} f_1(p, t) dp = \frac{\alpha}{2} = \int_{x_2(t)}^1 f_1(p, t) dp.$$

Further relevant information that can be determined from the 1-p.d.f. includes:

Distribution of time until a given proportion of susceptibles remains in the population

This distribution answers the following question:

What is the expected time before $\rho = 80\%$ of the population remains susceptible?

This distribution is computed from the solution of the SI-model:

$$P(T) = \frac{P_0}{e^{\beta T}(1-P_0)+P_0} \Rightarrow \{\rho = P(T)\} \Rightarrow T = \frac{1}{\beta} \ln \left(\frac{P_0(1-\rho)}{\rho(1-P_0)} \right).$$

Now, we apply two-dimensional R.V.T. technique to

$$\begin{aligned} X_1 &= \beta, & Y_1 &= T, & Y_1 = r_1(X_1, X_2) &= \frac{\ln(\frac{X_2(1-\rho)}{\rho(1-X_2)})}{X_1}, & X_1 = s_1(Y_1, Y_2) &= \frac{\ln(\frac{Y_2(1-\rho)}{\rho(1-Y_2)})}{Y_1}, \\ X_2 &= P_0, & Y_2 &= P_0, & Y_2 = r_2(X_1, X_2) &= X_2, & X_2 = s_2(Y_1, Y_2) &= Y_2, \end{aligned}$$

and taking into account that $\frac{\partial s_2(y_1, y_2)}{\partial y_1} = 0$, the jacobian is

$$J = -\frac{1}{(y_1)^2} \ln \left(\frac{y_2(1-\rho)}{\rho(1-y_2)} \right) \neq 0,$$

hence the joint p.d.f. of $(Y_1, Y_2) = (T, P_0)$ is given by

$$\begin{aligned} f_{T,P_0}(t, p_0) &= \frac{1}{t^2} \left| \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) \right| f_{\beta, P_0} \left(\frac{1}{t} \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right), p_0 \right) \\ &= \frac{1}{t^2} \left| \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) \right| f_\beta \left(\frac{1}{t} \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) \right) f_{P_0}(p_0), \end{aligned}$$

Therefore, the P_0 -marginal distribution of $f_{T,P_0}(t, p_0)$ yields the p.d.f. of T

$$f_T(t; \rho) = \frac{1}{t^2} \int_{\max(p_{0,1}, c_1)}^{\min(p_{0,2}, c_2)} \left| \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) \right| f_\beta \left(\frac{1}{t} \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) \right) f_{P_0}(p_0) dp_0, \quad p_0 \in D_{P_0},$$

where

$$D_{P_0} = \{p_0 = P_0(\omega), \omega \in \Omega : 0 \leq p_{0,1} \leq p_0 \leq p_{0,2} \leq 1\}$$

and

$$c_1 = \frac{\rho e^{\beta_1 t}}{\rho e^{\beta_1 t} + (1-\rho)}, \quad c_2 = \frac{\rho e^{\beta_2 t}}{\rho e^{\beta_2 t} + (1-\rho)}.$$

For t and ρ previously fixed, these values have been determined by imposing that

$$\beta_1 < \frac{1}{t} \ln \left(\frac{p_0(1-\rho)}{\rho(1-p_0)} \right) < \beta_2,$$

being

$$D_\beta = \{\beta = \beta(\omega), \omega \in \Omega : 0 \leq \beta_1 \leq \beta \leq \beta_2 \leq 1\}.$$

Modelling the diffusion of a new technology

year	1995	1996	1997	1998	1999	2000	2001	2002	2003
penetration rate (x_i)	2.3	7.5	10.2	16.2	37.3	59.9	72.6	81.9	89.3
year	2004	2005	2006	2007	2008	2009	2010	2011	---
penetration rate (x_i)	91.2	99.2	104.4	108.9	109.6	111.4	111.7	113.9	---

Remarks:

- x_i represents the rate of mobile phone lines per 100 inhabitants taking as reference the Spanish census corresponding to year 2011 updated by INE (National Statistics Institute of Spain).
- x_i , may be greater than 100% since any individual can possess more than one mobile phone line. In order to be able to apply the SI-model, two transformations on the data listed in previous table will be done.

Transformation:

$$P_i = 1 - x_i / 115, \quad i = 0, 1, \dots, 16 \Rightarrow 0 \leq P_i \leq 1.$$

- ① Standardize the values x_i by assuming a saturation value of 115.
- ② Since the unknown $P(t)$ of SI-model represents the percentage of susceptibles instead of infected (i.e., the percentage of people who have already adopted the mobile phone technology).

year	1995	1996	1997	1998	1999	2000	2001	2002	2003
P_i	0.9800	0.9348	0.9113	0.8591	0.6757	0.4791	0.3687	0.2878	0.2235
year	2004	2005	2006	2007	2008	2009	2010	2011	---
P_i	0.2070	0.1374	0.0922	0.0530	0.0470	0.0313	0.028695	0.0096	---

Assumptions:

$$P_i \in (0, 1) \Rightarrow P_0 \sim \text{Be}(a; b), \quad \beta > 0 \Rightarrow \beta \sim \text{Ga}(\lambda; \tau).$$

Fitting the model parameters: Determining a, b, λ, τ :

- ① Split the sample data: We take data from $t_0 = 1995$ to $t_{12} = 2007$.
- ② Minimizing the mean square error:

$$\min_{a, b, \lambda, \tau > 0} E(a, b, \lambda, \tau) = \sum_{i=0}^{12} (P_i - \mathbb{E}[P(t; a, b, \lambda, \tau)])^2 = \sum_{i=0}^{12} \left(P_i - \int_0^1 p f_1(p, t) dp \right)^2$$

where

$$f_1(p, t) = \frac{1}{t} \int_1^\infty f_{P_0} \left(\frac{p\xi}{1-p+p\xi} \right) f_\beta \left(\frac{\ln(\xi)}{t} \right) \frac{1}{(1-p+p\xi)^2} d\xi,$$

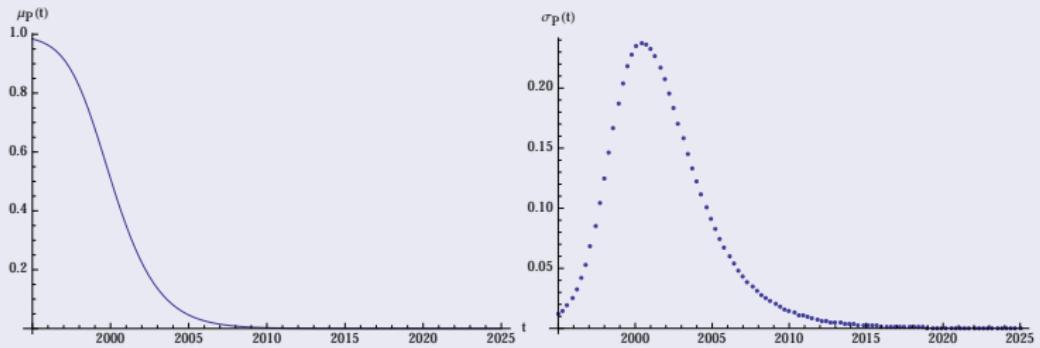
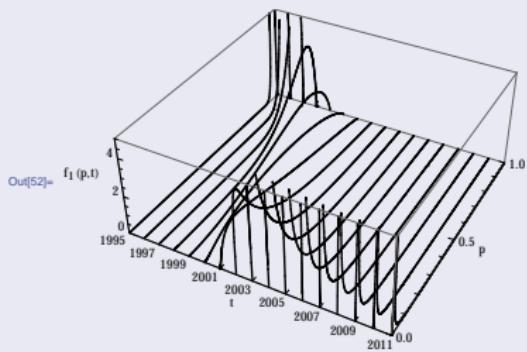
$$f_{P_0} \left(\frac{p\xi}{1-p+p\xi} \right) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{p\xi}{1-p+p\xi} \right)^{a-1} \left(\frac{1-p}{1-p+p\xi} \right)^{b-1},$$

$$f_\beta \left(\frac{\ln(\xi)}{t} \right) = \lambda^{\tau\xi - \frac{\lambda}{t}} \frac{1}{\Gamma(\tau)} \left(\frac{\ln(\xi)}{t} \right)^{\tau-1}.$$

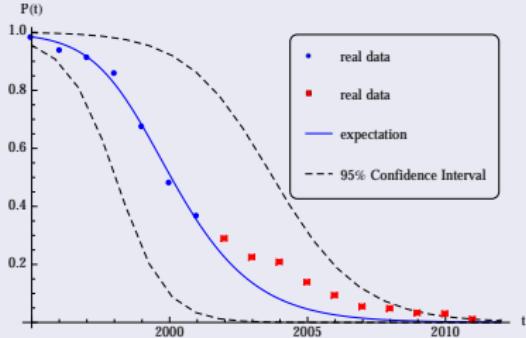


Using the Nelder-Mead algorithm we obtain:

$$a^* = 114.95, \quad b^* = 1.83, \quad \lambda^* = 27.36, \quad \tau^* = 0.032.$$

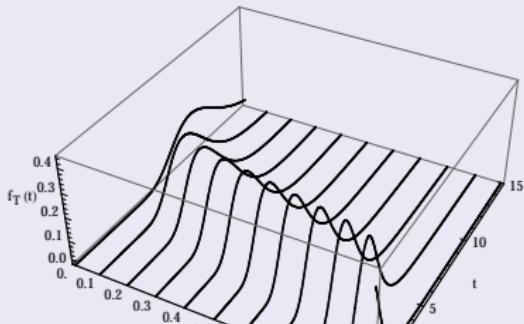


Validation and prediction using confidence intervals

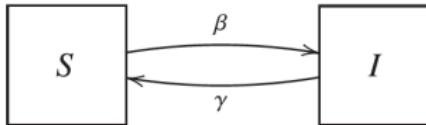


P.d.f. of the time T until a proportion $\rho = 90\%$ of susceptibles remain in the population

$$\mathbb{E}[T] = \int_0^{\infty} t f_T(t; 0.90) dt = 2.65$$



Motivating the nonlinear case: The SIS-type epidemiological model



- $S(t)$ number of susceptibles in the time instant t .
- $I(t)$ number of infected in the time instant t .
- n size of the total population. It is assumed to be constant for all time t .
- $\beta > 0$ rate of decline in the number of susceptibles.
- $\gamma > 0$ rate of infected that recover from the disease.

$$\begin{cases} S'(t) = -\beta S(t)I(t) + \gamma I(t), \\ I'(t) = \beta S(t)I(t) - \gamma I(t), \end{cases} \quad t > 0, \quad S(0) = S_0, \quad I(0) = I_0,$$

Taking into account that the solution $(S(t), I(t))$ of the SIS-model can be written as follows

$$\begin{aligned} S(t) &= \frac{\gamma(1 - S_0) + (S_0\beta - \gamma)e^{(\gamma-\beta)t}}{\beta(1 - S_0) + (S_0\beta - \gamma)e^{(\gamma-\beta)t}}, & t \geq 0. \\ I(t) &= \frac{(\beta - \gamma)(1 - S_0)e^{(\beta-\gamma)t}}{\beta(1 - S_0)e^{(\beta-\gamma)t} + S_0\beta - \gamma}, \end{aligned} \quad (2)$$

Applying the RVT technique, one can establish the following results

1-p.d.f. of the solution s.p. of the SIS-type epidemiological model

$$f_1(s, t) = \int_{\mathcal{D}_\gamma} \int_{\mathcal{D}_\beta} f_{S_0, \gamma, \beta} \left(\frac{\xi + e^{(\xi-\eta)t}(-1+s)\xi - s\eta}{\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta}, \xi, \eta \right) \frac{e^{(\xi-\eta)t}(\xi - \eta)^2 d\eta d\xi}{(\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta)^2},$$

$$f_1(i, t) = \int_{\mathcal{D}_\gamma} \int_{\mathcal{D}_\beta} f_{S_0, \gamma, \beta} \left(\frac{\xi - \eta - e^{(\xi-\eta)t}i\xi + i\eta}{\xi - \eta - e^{(\xi-\eta)t}i\eta + i\eta}, \xi, \eta \right) \frac{e^{(\xi-\eta)t}(\xi - \eta)^2 d\eta d\xi}{(\xi - \eta - e^{(\xi-\eta)t}i\eta + i\eta)^2},$$

where

$$\mathcal{D}_\beta, \quad \mathcal{D}_\gamma,$$

are the domains of r.v.'s β and γ , respectively.

Distribution of time until a given proportion of susceptibles and infected remains in the population

$$\begin{aligned}f_1(t, \rho_S) &= \int_{\mathcal{D}_\gamma} \int_{\mathcal{D}_\beta} f_{S_0, \gamma, \beta} \left(\frac{\xi(1 + e^{t(\xi-\eta)}(-1 + \rho_S)) - \eta\rho_S}{\xi + \eta(e^{t(\xi-\eta)}(-1 + \rho_S) - \rho_S)}, \xi, \eta \right) \\&\quad \times \frac{e^{t(\xi-\eta)}(\xi - \eta)^2(1 - \rho_S)|\xi - \eta\rho_S|}{(\xi + \eta(e^{t(\xi-\eta)}(-1 + \rho_S) - \rho_S))^2} d\eta d\xi. \\f_1(t, \rho_I) &= \int_{\mathcal{D}_\gamma} \int_{\mathcal{D}_\beta} f_{S_0, \gamma, \beta} \left(\frac{\xi + \eta(-1 + \rho_I) - e^{t(\xi-\eta)}}{\xi - \eta(1 + (-1 + e^{t(\xi-\eta)})\rho_I)}, \xi, \eta \right) \\&\quad \times \left| \frac{e^{t(\xi-\eta)}(\xi - \eta)^2(\xi + \eta(-1 + \rho_I))\rho_I}{(\xi - \eta(1 + (-1 + e^{t(\xi-\eta)})\rho_I))^2} \right| d\eta d\xi.\end{aligned}$$

In epidemiology, the **basic reproduction number**, R_0 , associated to an infection is useful to elucidate whether will spread out or not. In the case of the SIS model, this value and its relationship with the propagation of the epidemic in the long run is given by

$$R_0 = \frac{\beta}{\gamma}, \quad \begin{cases} \text{if } R_0 < 1 \equiv \beta < \gamma, & \text{then the diseases will die out as } t \rightarrow +\infty, \\ \text{if } R_0 > 1 \equiv \beta > \gamma, & \text{then the diseases will spread out as } t \rightarrow +\infty. \end{cases}$$

This classification is easily derived from expression of $I(t)$, or equivalently of $S(t)$, since

$$\lim_{t \rightarrow +\infty} I(t) = \lim_{t \rightarrow +\infty} \frac{(\beta - \gamma)(1 - S_0)e^{(\beta - \gamma)t}}{\beta(1 - S_0)e^{(\beta - \gamma)t} + S_0\beta - \gamma} = 0 \quad \text{if } \beta < \gamma,$$

$$\lim_{t \rightarrow +\infty} S(t) = \lim_{t \rightarrow +\infty} \frac{\gamma(1 - S_0) + (S_0\beta - \gamma)e^{(\gamma - \beta)t}}{\beta(1 - S_0) + (S_0\beta - \gamma)e^{(\gamma - \beta)t}} = 1 \quad \text{if } \beta < \gamma.$$

In our context, both β and γ are r.v.'s, so that the requirement for epidemic extinction in the deterministic framework $\beta < \gamma$ means the computation of the following probability in the stochastic scenario

$$\mathbb{P}[\mathcal{S}], \quad \mathcal{S} = \{\omega \in \Omega : \beta(\omega) < \gamma(\omega)\} = \{\omega \in \Omega : R_0(\omega) < 1\}. \quad (3)$$

This key probability can be computed by taking advantage of RVT. Using the mapping

$$\mathbf{U} = (U_1, U_2)^T = (\gamma, \beta)^T, \quad V = \frac{U_2}{U_1} = \frac{\beta}{\gamma} = R_0,$$

one gets

$$f_{R_0}(r_0) = \int_{\mathcal{D}(\gamma)} f_{\gamma, \beta}(\xi, r_0 \xi) |\xi| d\xi,$$

where $f_{\gamma, \beta}(\cdot, \cdot)$ denotes the (γ, β) -marginal distribution of the joint p.d.f. of the random inputs (S_0, γ, β) . This allows us to compute the target probability

$$\mathbb{P}[\mathcal{S}] = \int_0^1 \int_{\mathcal{D}(\gamma)} f_{\gamma, \beta}(\xi, r_0 \xi) |\xi| d\xi dr_0.$$

Modelling the spread of smoking in Spain

year (t_j)	1987 ($j = 0$)	1993 ($j = 6$)	1995 ($j = 8$)	1997 ($j = 10$)	2001 ($j = 14$)	2003 ($j = 16$)	2006 ($j = 19$)
S_j	0.4488	0.5144	0.5278	0.5514	0.5783	0.6244	0.6467

$$\mathcal{J} = \{0, 6, 8, 10, 14, 16, 19\}$$

Assumptions:

$$S_0 \sim \text{Be}(a; b); \quad \beta > 0 \Rightarrow \beta \sim \text{Exp}_{[0, 1000]}(\lambda_\beta); \quad \gamma > 0 \Rightarrow \gamma \sim N_{[0, 1]}(\mu_\gamma; \sigma_\gamma).$$

Fitting the model parameters: Determining $a, b, \lambda_\beta, \mu_\gamma, \sigma_\gamma$:

$$\min_{a, b, \lambda_\beta, \mu_\gamma, \sigma_\gamma > 0} E(a, b, \lambda_\beta, \mu_\gamma, \sigma_\gamma) = \sum_{j \in \mathcal{J}} (S_j - \mathbb{E}[S(t_j; a, b, \lambda_\beta, \mu_\gamma, \sigma_\gamma)])^2,$$

where,

$$\mathbb{E}[S(t_j; a, b, \lambda_\beta, \mu_\gamma, \sigma_\gamma)] = \int_0^1 s f_1(s, t_j) ds, \quad j \in \mathcal{J}.$$

$$\begin{aligned} f_1(s, t) &= \int_0^1 \int_0^{1000} f_{S_0} \left(\frac{\xi + e^{(\xi-\eta)t}(-1+s)\xi - s\eta}{\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta} \right) f_\gamma(\xi) f_\beta(\eta) \\ &\times \frac{e^{(\xi-\eta)t}(\xi - \eta)^2 d\eta d\xi}{(\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta)^2}. \end{aligned}$$

Note that the p.d.f.'s for input data are

$$f_{S_0} \left(\frac{\xi + e^{(\xi-\eta)t}(-1+s)\xi - s\eta}{\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta} \right) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{\xi + e^{(\xi-\eta)t}(-1+s)\xi - s\eta}{\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta} \right)^{a-1} \\ \times \left(\frac{e^{(\xi-\eta)t}(-1+s)(\eta - \xi)}{\xi + e^{(\xi-\eta)t}(-1+s)\eta - s\eta} \right)^{b-1},$$

$$f_\beta(\eta) = \frac{\lambda_\beta e^{-\lambda_\beta \eta}}{\int_0^{1000} \lambda_\beta e^{-\lambda_\beta \eta} d\eta},$$

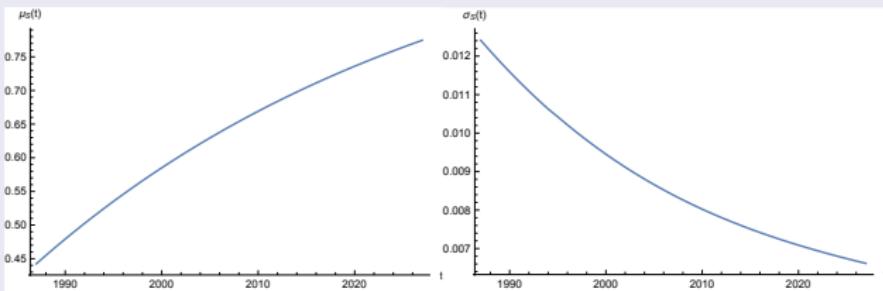
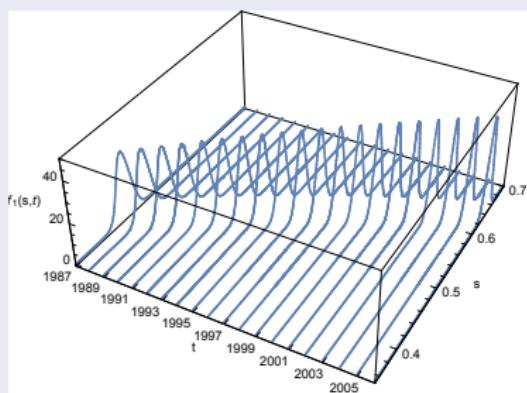
and

$$f_\gamma(\xi) = \begin{cases} e^{-\frac{(\xi-\mu_\gamma)^2}{2(\sigma_\gamma)^2}} \frac{1}{\sqrt{2\pi}\sigma_\gamma \left(\frac{1}{2} \operatorname{erfc} \left(\frac{\mu_\gamma-1}{\sqrt{2}\sigma_\gamma} \right) - \frac{1}{2} \operatorname{erfc} \left(\frac{\mu_\gamma}{\sqrt{2}\sigma_\gamma} \right) \right)}, & \text{if } 0 < \xi \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

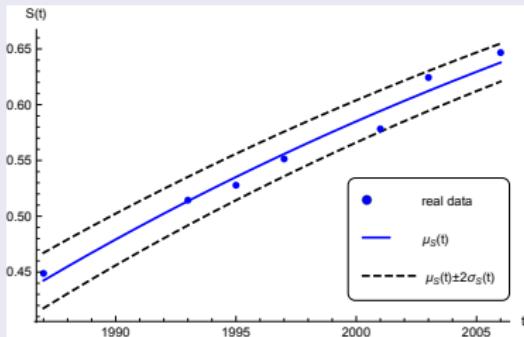
being $\operatorname{erfc}(z) = 1 - \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ the complementary error function.

Using the Nelder-Mead algorithm we obtain:

$$a^* = 708.755, \quad b^* = 893.394, \quad \lambda_\beta^* = 1362.230, \quad \mu_\gamma^* = 0.0231162, \quad \sigma_\gamma^* = 0.0000526.$$



Validation and prediction using confidence intervals



year (t_j)	1987 ($j = 0$)	1993 ($j = 6$)	1995 ($j = 8$)	1997 ($j = 10$)	2001 ($j = 14$)	2003 ($j = 16$)	2006 ($j = 19$)
Confidence level	0.9550	0.9544	0.9545	0.9546	0.9549	0.9550	0.9552

Table: Probabilities associated to the confidence intervals built according to the SIS model.

Expected time until a certain proportion ρ_S of the population remains non-smoker

$$\begin{aligned}f_1(t, \rho_S) &= \int_0^1 \int_0^{+\infty} f_{S_0} \left(\frac{\xi(1+e^{t(\xi-\eta)}(-1+\rho_S)) - \eta\rho_S}{\xi + \eta(e^{t(\xi-\eta)}(-1+\rho_S) - \rho_S)} \right) f_\gamma(\xi) f_\beta(\eta) \\&\times \frac{e^{t(\xi-\eta)}(\xi-\eta)^2(1-\rho_S)|\xi-\eta\rho_S|}{(\xi + \eta(e^{t(\xi-\eta)}(-1+\rho_S) - \rho_S))^2} d\eta d\xi,\end{aligned}$$

ρ_S	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85
$\mathbb{E}[T_S]$	0.59	4.78	9.42	14.61	20.51	27.32	35.40	45.30	58.10

Table: Expectation of time T_S until a proportion, ρ_S , of the population remains non-smoker for different values ρ_S .

$$\mathbb{E}[T_S] = \int_0^\infty t f_{T_S}(t; 0.75) dt = 35.4013.$$

This means that the middle of the year 2023 approximately represents the average time until 75% of the Spanish men aged over 16 years old population will be non-smokers.

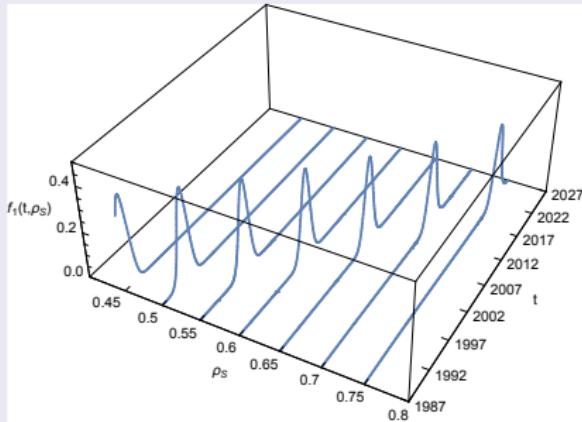


Figure: Plot of the 1-p.d.f. of the time T_S until a proportion $p_S \in \{0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75\}$ of the population remains susceptible.

Finally, we compute the probability of the event \mathcal{S} previously introduced

$$\mathbb{P}[\mathcal{S}] = \int_0^1 \int_0^1 f_\gamma(\xi) f_\beta(r_0 \xi) |\xi| dr_0 d\xi = 0.999453,$$

where $f_\beta(\eta)$ and $f_\gamma(\xi)$ are the p.d.f.'s of β and γ , respectively. This is in accordance with the interpretation of the basic reproductive number R_0 : the percentage of Spanish smoker men older than 16 years old will likely disappear as t tends to $+\infty$.

The randomized Bertalanffy model: A fish weight growth over the time.

$$\begin{aligned}\dot{W}(t) &= -\lambda W(t) + \eta(W(t))^{2/3}, & t \geq t_0, \\ W(t_0) &= W_0.\end{aligned}\quad \left. \right\}$$

- $W(t)$: fish weight growth at time instant t .
- η : intrinsic growth rate.
- λ : linear coefficient.
- We shall assume that all these inputs are r.v.'s with joint p.d.f. $f_{W_0, \eta, \lambda}(w_0, \eta, \lambda)$ and

$$\mathbb{P}[\{\omega \in \Omega : W_0(\omega) \neq 0\}] = 1, \quad \mathbb{P}[\{\omega \in \Omega : \lambda(\omega) \neq 0\}] = 1.$$

Our goals are

1. To determine the 1-p.d.f. of the solution applying RVT method.
2. To use real data in order to assign a reliable probabilistic distribution to random inputs using an **inverse frequentist technique**.
3. To construct both punctual and probabilistic predictions based on confidence intervals.

Step 1: To determine the 1-p.d.f. of the solution

Considering the change of variable $W(t) = (Z(t))^3$, applying the following result

A particular case of RVT

Let $Z : \Omega \rightarrow \mathbb{R}$ be an absolutely continuous real r.v. defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, with p.d.f. $f_Z(z)$. Assume that $Z(\omega) \neq 0$ for all $\omega \in \Omega$. Then, the p.d.f. $f_W(w)$ of the transformation $W = Z^3$ is given by

$$f_W(w) = \frac{1}{3} f_Z(\sqrt[3]{w}) |w|^{-2/3}.$$

Then, the 1-p.d.f. of the solution $W(t)$ of the Bertalanffy model is

$$\begin{aligned} f_1(w, t) &= \frac{1}{3} f_Z(w^{1/3}) |w|^{-2/3} \\ &= \int_{\mathcal{D}(\eta)} \int_{\mathcal{D}(\lambda)} f_{W_0, \eta, \lambda} \left(\left(\frac{e^{(1/3)\lambda(t-t_0)} \lambda w^{1/3} + \eta - e^{(1/3)\lambda(t-t_0)} \eta}{\lambda} \right)^3 \eta \lambda \right) \\ &\quad \times \left(\frac{e^{(1/3)\lambda(t-t_0)} \lambda w^{1/3} + \eta - e^{(1/3)\lambda(t-t_0)} \eta}{\lambda} \right)^2 e^{(1/3)\lambda(t-t_0)} |w|^{-2/3} d\lambda d\eta. \end{aligned}$$

Step 2: To assign a probabilistic distribution to random inputs using real data

t_i (years)	1	2	3	4	5	6	7
w_i (lbs)	0.2	0.4	0.6	0.9	1	1.3	1.6
t_i (years)	8	9	10	11	12	13	14
w_i (lbs)	1.8	2.3	2.6	2.9	3.1	3.4	3.7
t_i (years)	15	16	17	18	19	20	21
w_i (lbs)	4.5	5.2	5.7	6.2	6.5	6.7	6.8
t_i (years)	22	23	24	25	26	27	28
w_i (lbs)	7.2	8.2	9	9.5	10	10.5	11
t_i (years)	29	30	31	32	33		
w_i (lbs)	11.5	12	12.5	13	14		

Table: Fish weights w_i for walleye species in lbs every year t_i , $1 \leq i \leq 33 = N$.

To assign a reliable probability distribution to these data, **Frequentist Inverse Technique** will be applied.

STEP 2.1: It is assumed that the **measured quantity of interest**, fish weights (w_i), are corrupted by measurement errors ε_i .

$$w_i = W(t_i; \mathbf{q}) = W(t_i; w_0, \eta, \lambda) + \varepsilon_i, \quad 1 \leq i \leq 33 = N,$$

where errors are assumed i.i.d. and $\varepsilon_i \sim N(0; \sigma^2)$, being $\sigma > 0$ fixed but unknown. As a consequence of this assignment the probabilistic distribution for the random vector $\mathbf{Q} = (W_0, \eta, \lambda)$ is assumed to be

$$\mathbf{Q} = (W_0, \eta, \lambda) \sim N_3(\mu_{\mathbf{Q}}; \Sigma_{\mathbf{Q}}),$$

where

- $\mu_{\mathbf{Q}} = (\hat{w}_0, \hat{\eta}, \hat{\lambda})$ is defined from appropriate estimates of (w_0, η, λ) .
- $\Sigma_{\mathbf{Q}}$ represents the variance-covariance matrix.

STEP 2.2: A least squares fit to the data yields the following parameter estimates $\mu_Q = (\hat{w}_0, \hat{\eta}, \hat{\lambda})$

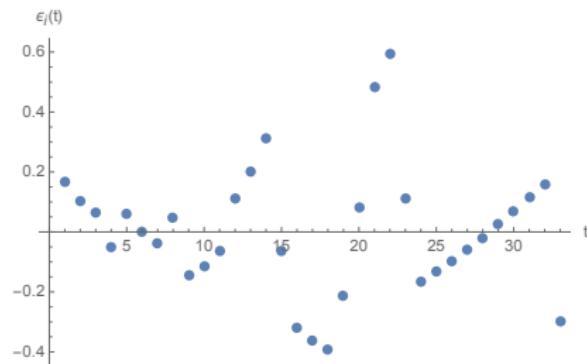
$$\hat{w}_0 = 0.365934, \quad \hat{\eta} = 0.305461, \quad \hat{\lambda} = 0.0880184.$$

The residuals of the fitting are,

$$\varepsilon_i = W(t_i; \hat{w}_0, \hat{\eta}, \hat{\lambda}) - w_i, \quad 1 \leq i \leq 33 = N.$$

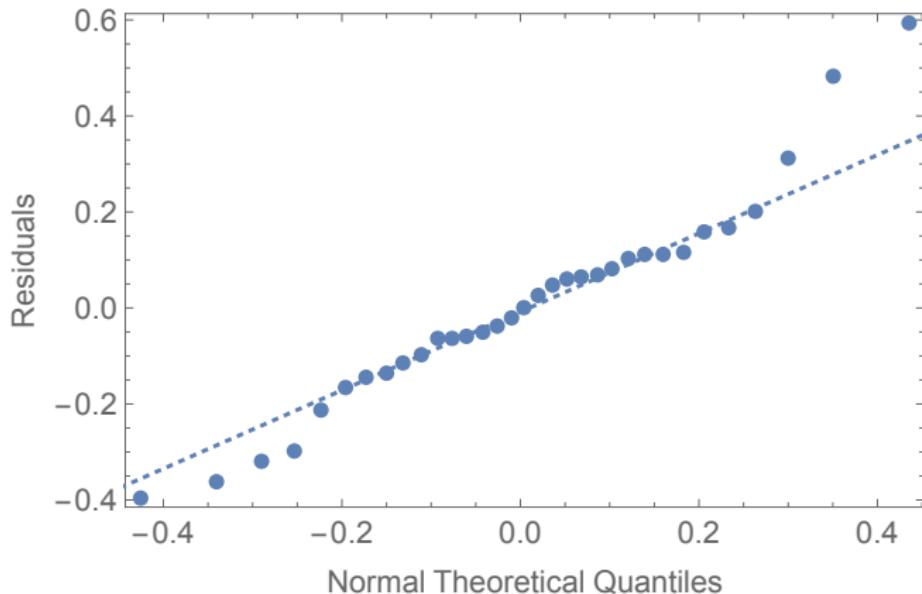
We need to check

- i.i.d.



- normality

Normality Test	Statistic	p-value
Shapiro-Walk Test	0.958995	0.242077



STEP 2.3: Determine the sensibility matrix

According to frequentist parameter estimation method first we compute the sensitivity matrix

$$\chi(\mathbf{Q}) = \left[\begin{array}{ccc} \frac{\partial W(t_1; \mathbf{Q})}{\partial W_0} & \dots & \frac{\partial W(t_{33}; \mathbf{Q})}{\partial W_0} \\ \frac{\partial W(t_1; \mathbf{Q})}{\partial \eta} & \dots & \frac{\partial W(t_{33}; \mathbf{Q})}{\partial \eta} \\ \frac{\partial W(t_1; \mathbf{Q})}{\partial \lambda} & \dots & \frac{\partial W(t_{33}; \mathbf{Q})}{\partial \lambda} \end{array} \right]^T \Big|_{\mathbf{Q}=(\hat{w}_0, \hat{\eta}, \hat{\lambda})}.$$

from the solution

$$W(t) = (W_0)^{1/3} e^{-(1/3)\lambda(t-t_0)} - \frac{\eta}{\lambda} e^{-(1/3)\lambda(t-t_0)} + \frac{\eta}{\lambda}.$$

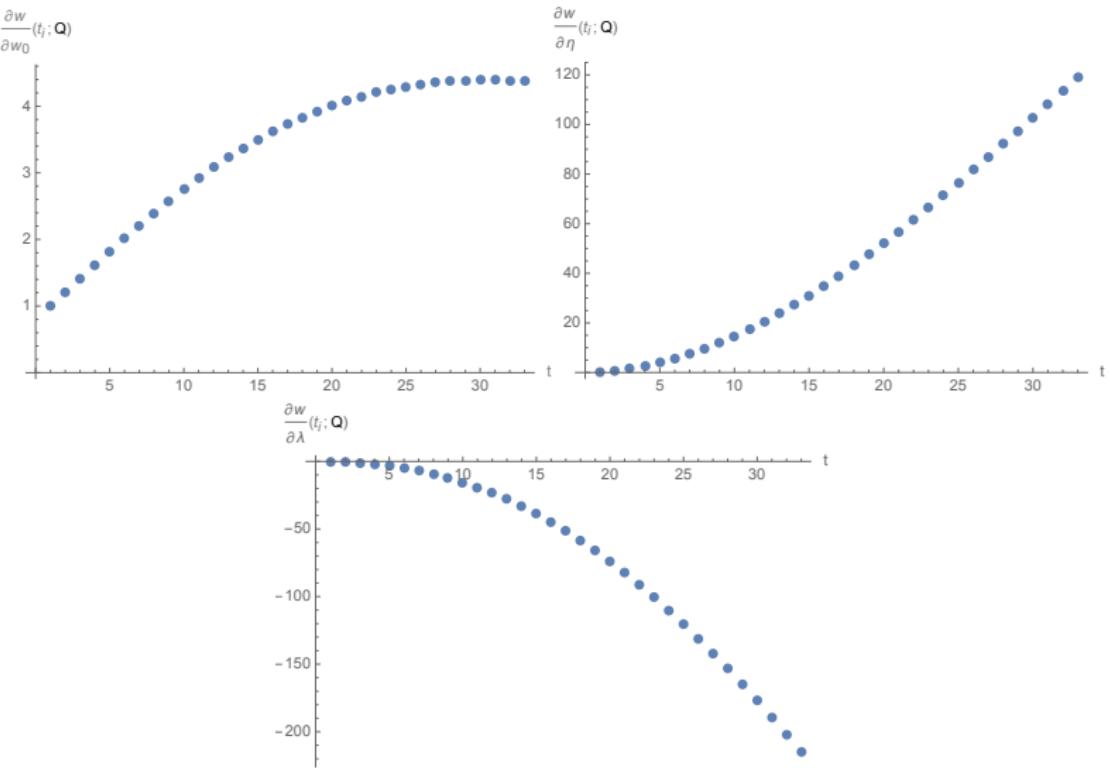


Figure: Top: Left: $\frac{\partial W}{\partial w_0}(t_i; \mathbf{Q})$. Right: $\frac{\partial W}{\partial \eta}(t_i; \mathbf{Q})$. Bottom: $\frac{\partial W}{\partial \lambda}(t_i; \mathbf{Q})$.
 $t_i = i, 1 \leq i \leq 33 = N$.

Then, the following probabilistic distribution has been assigned to model parameters

$$\mathbf{Q} = (W_0, \eta, \lambda) \sim N_3(\mu_{\mathbf{Q}}; \Sigma_{\mathbf{Q}}),$$

where

- The mean vector has been previously estimated by mean square method

$$\mu_{\mathbf{Q}} = (0.365934, 0.305461, 0.0880184).$$

- The variance-covariance matrix is

$$\Sigma_{\mathbf{Q}} = \sigma^2 \left((\chi(\mathbf{Q}))^T \chi(\mathbf{Q}) \right)^{-1} = \begin{bmatrix} 0.0029288 & -0.000812275 & -0.000400288 \\ -0.00081227 & 0.000268075 & 0.000136915 \\ -0.000400288 & 0.000136915 & 0.0000705259 \end{bmatrix},$$

being σ the error standard deviation estimate:

$$\sigma = \sqrt{\sum_{i=1}^{33} (\varepsilon_i)^2} = 0.214435.$$

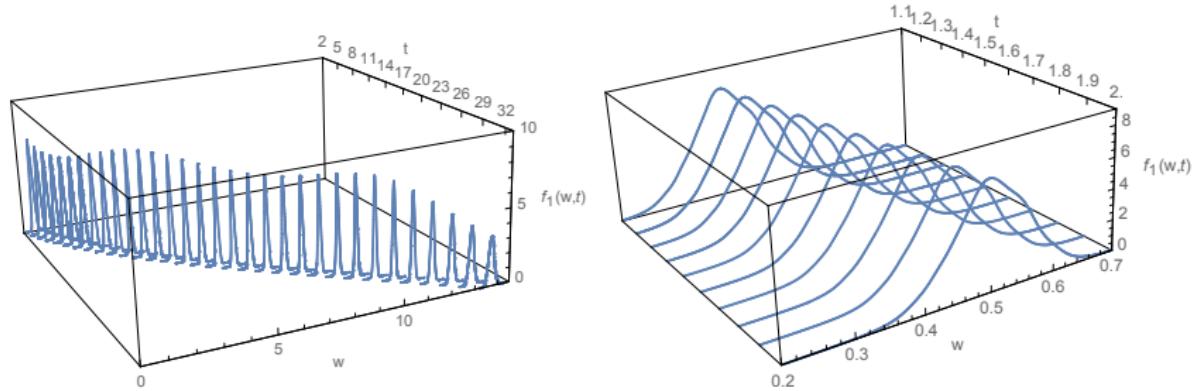


Figure: Left: 1-p.d.f. of the solution stochastic process to random Bertalanffy model given for all the times of the sample, $t \in \{2, \dots, 33 = N\}$. Right: Detailed representation of the 1-p.d.f. for the times $t \in \{1.1, 1.2, \dots, 2\}$.

Step 3: To construct punctual and probabilistic predictions based on confidence intervals

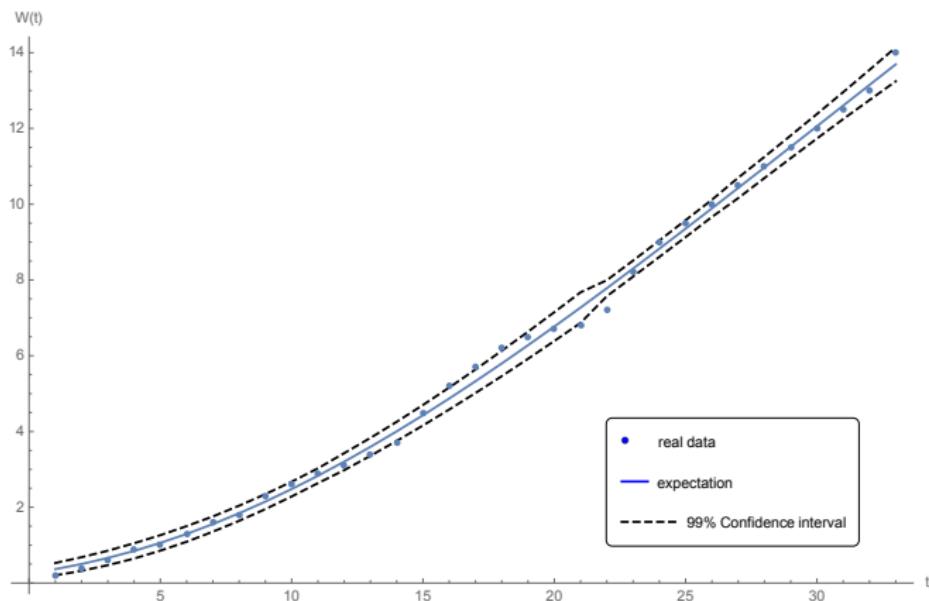


Figure: Expectation (solid line) and 99%-confidence intervals (dotted lines). Points represent fish weigh.

Conclusions and forthcoming work

- ① Random Variable Transformation (R.V.T.) method is a powerful tool to compute the 1-p.d.f. of the solution stochastic process of Random Differential Equations.
- ② The application of this technique has been shown for first order lineal and nonlinear random differential equations, but it can be extended to second-order differential equations and random difference equations.
- ③ Extension of the results for systems of both random differential and difference equations.
- ④ Application of R.V.T. technique together with numerical methods.

References (Continuous models based on Random Differential Equations)

- ① Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2014): Determining the first probability density function of linear random initial value problems by the Random Variable Transformation (R.V.T.) technique: A comprehensive study, *Abstr. & Appl. Anal.* 2014- ID248512, pp: 1–25.
- ② Casabán M.C., Cortés J.C., Navarro-Quilés, A., Romero J.V., Roselló M.D., Villanueva, R.J. (2015): Probabilistic solution of the random homogeneous Riccati differential equation: A comprehensive case-study by using linearization and the Random Variable Transformation techniques, *J. Comput. Appl. Math.* 24, pp: 20–35.
- ③ Casabán M.C., Cortés J.C., Navarro-Quilés, A., Romero J.V., Roselló M.D., Villanueva, R.J. (2016): Computing probabilistic solutions of the Bernoulli random differential equation, *J. Comput. Appl. Math.* (accepted).
- ④ Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2016): Probabilistic solution of random autonomous first-order linear systems of ordinary differential equations, *Romanian Reports in Physics* (accepted).
- ⑤ Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2016): Solving random homogeneous linear second-order differential equations: A full probabilistic description, *Mediterranean J. of Mathematics* (accepted).

References (Discrete/Continuous Models based on Random Differential Equations)

- ① Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2016): Random first-order linear discrete models and their probabilistic solution: A comprehensive study, *Abstr. & Appl. Anal.* 2016-ID6372108, pp: 1–22.
- ② Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2014): Probabilistic solution of random homogeneous linear second-order difference equations, *Appl. Math. Lett.* 34, pp: 27–33.
- ③ Casabán M.C., Cortés J.C., Romero J.V., Roselló M.D. (2015): Probabilistic solution of random SI-type epidemiological models using the Random Variable Transformation technique, *Comm. Nonl. Sc. Num. Simul.* 24(1-3), pp: 86–97.
- ④ Casabán M.C., Cortés J.C., Navarro-Quilés, A., Romero J.V., Roselló M.D., Villanueva, R.J. (2016): A comprehensive probabilistic solution of random SIS-type epidemiological models using the Random Variable Transformation technique, *Comm. Nonl. Sc. Num. Simul.* 32, pp: 199–210.

Computational Methods for Random Epidemiological Models

Métodos Computacionales para el Estudio de Modelos Epidemiológicos con Incertidumbre

Conferencias de Investigación para Posgrado 2016
Universidad Complutense de Madrid

24 junio de 2016

Prof. Dr. Juan Carlos Cortés

Instituto Universitario de Matemática Multidisciplinar
Universitat Politècnica de València

*im*²



UNIVERSIDAD
POLITECNICA
DE VALENCIA