What is ecotoxicology?
Dose-response modelling
Parameter inference

Modelling of Biological Systems
- From the Individual to the Ecosystem -

Sandrine CHARLES (UCBL - LBBE)

5BS_BIM_OPTEcosy)

sandrine.charles@univ-lyon1.fr

Université Claude Bernard Lyon 1

What is ecotoxicology?
Dose-response modelling
Parameter inference

What is ecotoxicology? **Introduction**
Dose-response modelling **Monospecific toxicity tests**
Parameter inference **Critical Effect Concentrations**

Table of content

What is ecotoxicology?

Dose-response modelling

Parameter inference

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Detailed content

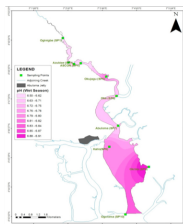| **What is ecotoxicology?** | **Introduction** |
| Dose-response modelling | Monospecific toxicity tests |
| Parameter inference | Critical Effect Concentrations |

Ecosystems are under...

...environmental pressure
due to variations in temperature, flow, conductivity...

https://climate.copernicus.eu/record-breaking-temperatures-june
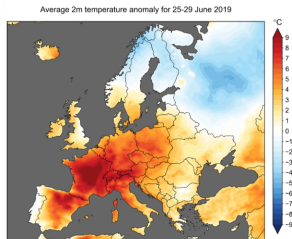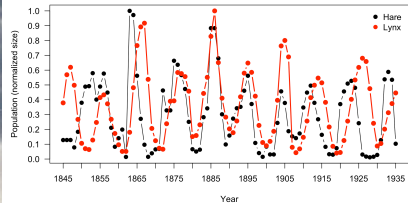
L.B. Ibanga, J.A. Nkwoji, A.I. Usese, I.C. Onyema, L.O. Chukwu. 2019. Hydrochemistry and heavy metals concentrations in sediment of Woji creek and Bonny estuary, Niger Delta, Nigeria. *Regional Studies in Marine Science*, 25, 10043.

https://www.eea.europa.eu/data-and-maps/figures/projected-change-in-damage-of-river-floods-with-a-100-year-return-period-between-2071-2100-and-1961-1990

What is ecotoxicology?	**Introduction**
Dose-response modelling	Monospecific toxicity tests
Parameter inference	Critical Effect Concentrations
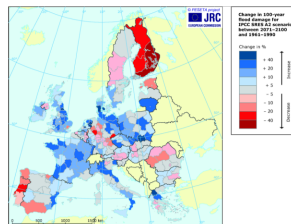
Ecosystems are under...

...ecological pressure
due to competition, predation, resource availability...

What is ecotoxicology? | **Introduction**
Dose-response modelling | Monospecific toxicity tests
Parameter inference | Critical Effect Concentrations

Ecosystems are under...

...chemical pressure
due to massive rejection of xenobiotics in air, soil and water

http://www.univers-
nature.com/actualite/le-
reste/plus-dune-commune-sur-
deux-concernees-par-la-
pollution-aux-nitrates-
57979.html

https://www.paperblog.fr/3707710/environnement-
la-nature-est-a-80-detruite-par-les-industries/

https://www.thehindu.com

What is ecotoxicology?	**Introduction**
Dose-response modelling	Monospecific toxicity tests
Parameter inference	Critical Effect Concentrations

Ecotoxicology

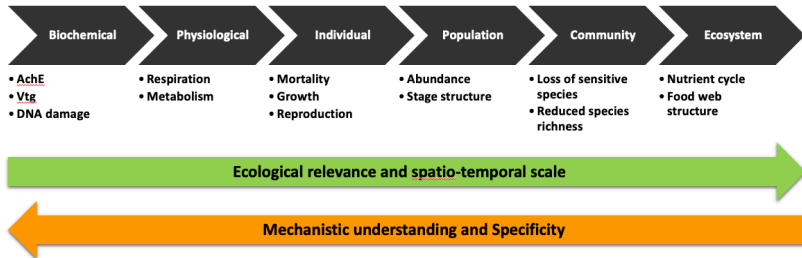▶ A scientific field at the bridge of chemistry, toxicology and ecology

The branch of toxicology concerned with the study of toxic effects, caused by natural or synthetic pollutants, to the constituents of ecosystems, animals (including humans), vegetables and microorganisms, in an integrated context [Truhaut, 1977]

Ecology in the presence of toxicants [Chapman, 2002]

▶ In ecotoxicology, the answer of the ecosystem to environmental perturbations (physical, chemical and/or biological) is studied in all compartments of the biosphere (air, soil and water) and at all levels of biological organization [Walker et al., 2006]

What is ecotoxicology?	**Introduction**
Dose-response modelling	Monospecific toxicity tests
Parameter inference	Critical Effect Concentrations

From one level of organization to the next

Depending on the level of biological organization, answers to chemical pressure may strongly differ [Clements, 2000]

What is ecotoxicology? **Introduction**
Dose-response modelling **Monospecific toxicity tests**
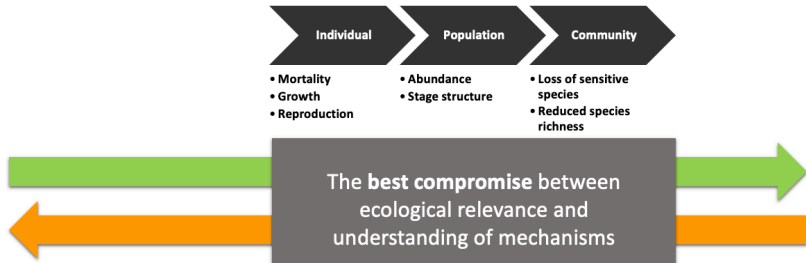**Parameter inference** **Critical Effect Concentrations**

From one level of organization to the next

Depending on the level of biological organization, answers to
chemical pressure may strongly differ [Clements, 2000]

What is ecotoxicology? | **Introduction**
Dose-response modelling | Monospecific toxicity tests
Parameter inference | Critical Effect Concentrations

A key challenge in ecotoxicology:
Extrapolating from one level to the next

- ▶ From the individual level...
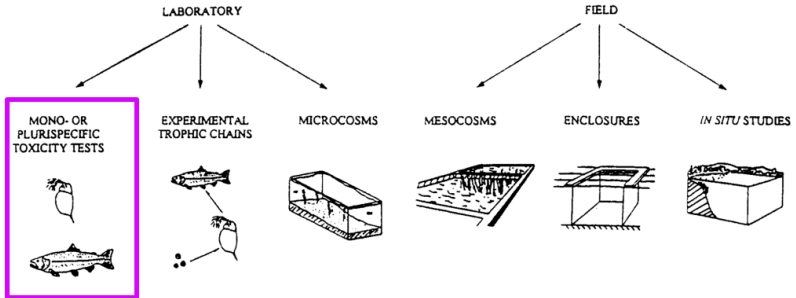 - Identify critical life history traits;
  - Identify chemical modes of action;
  - Account for time-dependent effects;
- ▶ ... to the population level...
  - Include individual effects in population dynamics;
  - Identify critical demographic parameters;
- ▶ ... to the community level
  - Species sensitivity distribution;
  - Community functioning accounting for ecological interactions.
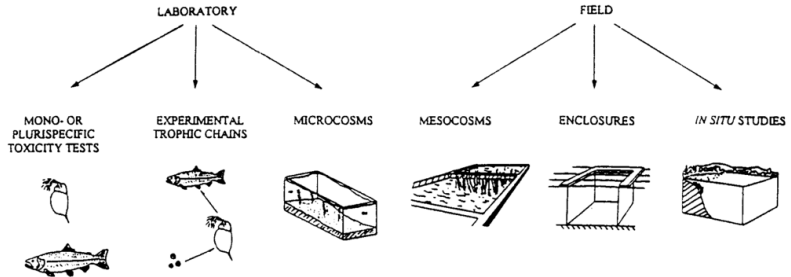
What is ecotoxicology?	**Introduction**
Dose-response modelling	Monospecific toxicity tests
Parameter inference	Critical Effect Concentrations

A variety of experimental devices

[Caquet et al., 1996]

What is ecotoxicology? **Introduction**
Dose-response modelling Monospecific toxicity tests
Parameter inference Critical Effect Concentrations

A variety of experimental devices

[Caquet et al., 1996]

What is ecotoxicology?	**Introduction**
Dose-response modelling	Monospecific toxicity tests
Parameter inference	Critical Effect Concentrations

Modelling at the individual level

▶ Estimating individual toxicity indices;

▶ Accounting for the individual variability.

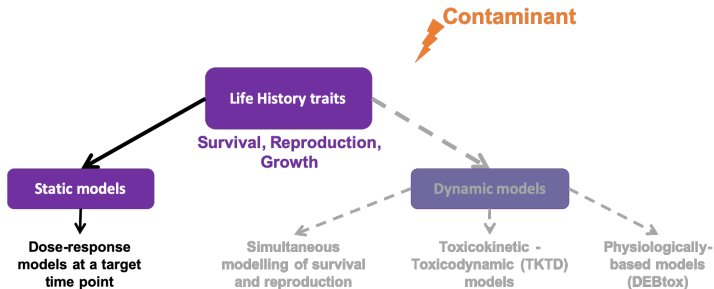| What is ecotoxicology? | **Introduction** |
| Dose-response modelling | Monospecific toxicity tests |
| Parameter inference | Critical Effect Concentrations |
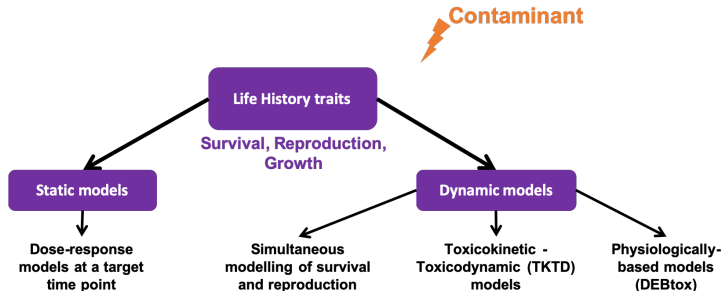
Modelling at the individual level

▶ Estimating individual toxicity indices;
▶ Accounting for the individual variability.

▶ Choose the appropriate model according to the data **and**
to the research question;
▶ Choose an inference method for estimating model
parameters.

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Detailed content

sandrine.charles@univ-lyon1.fr **BS-5-ECOSY**

What is ecotoxicology? | Introduction
Dose-response modelling | **Monospecific toxicity tests**
Parameter inference | Critical Effect Concentrations

Definition of toxicity tests or bioassays

► Bioassay is a word commonly used instead of **biological assay** or **toxicity tests**. It's a particular type of scientific experiment.

► Bioassays are typically conducted to measure the effects of potentially toxic substances on living organisms.

► Bioassays can be
  • **qualitative**: dedicated to assess physical effects of a substance that cannot be quantified (*e.g.*, abnormality or deformity).
  • **quantitative**: dedicated to estimate the potency of a substance by measurement of the biological response/effect it produces; quantitative bioassays are typically analyzed using **statistical methods**.

What is ecotoxicology?   Introduction
Dose-response modelling   **Monospecific toxicity tests**
Parameter inference   Critical Effect Concentrations

Definition of toxicity tests (*continued*)

Several kind of substances can be studied, for example:

- ▶ Pesticides, chemical, pharmaceutical, cosmetic substances;
- ▶ Effluents (industrial discharges, outputs from water plants);
- ▶ Polluted soils, waste, sewage sludge, sediments.

According to the substance, different kinds of experiments can be conducted:

- ▶ A control versus one treatment;
- ▶ A control and several treatments.

A treatment can be a fixed concentration of a substance (C_1, C_2...) or a time-variable concentration $C(t)$.
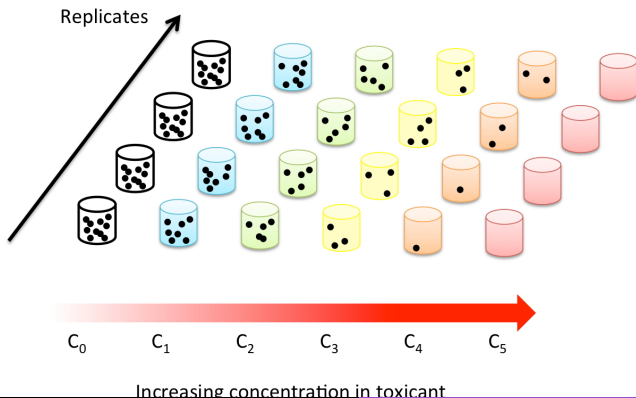
What is ecotoxicology?   Introduction
Dose-response modelling   **Monospecific toxicity tests**
Parameter inference   Critical Effect Concentrations

Acute vs. chronic quantitative bioassays

- ▶ **Acute toxicity**: from some hours to some days (*e.g.*, survival or mobility inhibition)
  - $\rightarrow$ short-time exposure at high concentrations;
  - $\rightarrow$ rapid impact on organisms.

- ▶ **Chronic toxicity**: from some days to some weeks (*e.g.*, growth or reproduction inhibition, sub-individual biomarkers)
  - $\rightarrow$ long-time exposure at low concentrations.

What is ecotoxicology?   Introduction
Dose-response modelling   **Monospecific toxicity tests**
Parameter inference   Critical Effect Concentrations

Standard experimental design

Under standardized protocols, individuals are counted **over time**, that is at regular time points
Endpoints can be survival, growth and/or reproduction for example.

Replicates

C_0 C_1 C_2 C_3 C_4 C_5

Increasing concentration in toxicant

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Example of a toxicity test

Daphnia magna, acute immobilisation test (OECD 202, 1984) and chronic reproduction test (OECD 211, 2012)
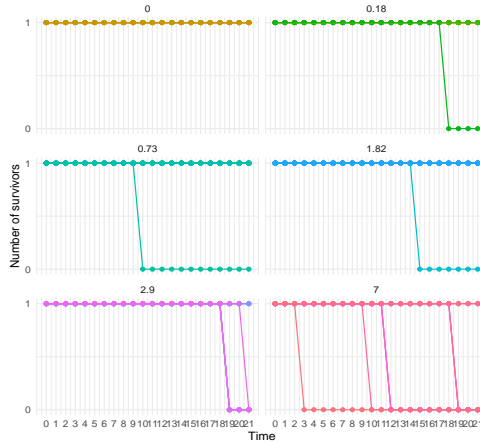

Daphnia magna

Acute test: the number of immobile daphnids is determined for each concentration at 24 and 48 hours.

Chronic test: offsprings are daily counted during 21 days.

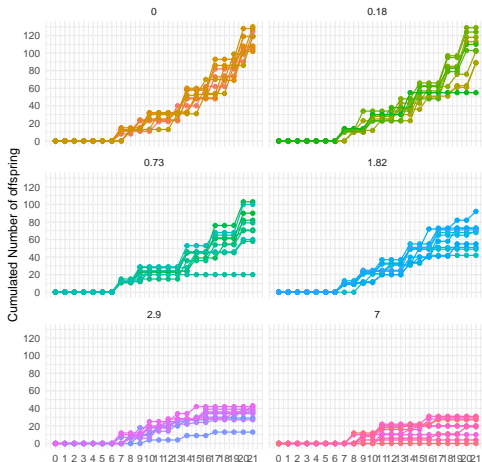| **What is ecotoxicology?** | **Introduction** |
| **Dose-response modelling** | **Monospecific toxicity tests** |
| **Parameter inference** | **Critical Effect Concentrations** |

Example of survival data

Effect of chlordane on *D. magna* survival during 21 days
(10 replicates of 1 individual):

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Example of reproduction data

Effect of chlordane on *D. magna* reproduction during 21 days (10 replicates of 1 individual):

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Detailed content

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Standard analyses of toxicity test data
Derive Critical Effect Concentrations

also called "summary statistics of toxicity" or "thresholds".

$\rightarrow$ Most common indicators to quantitatively assess risks for single species exposed to single contaminants.

$\rightarrow$ Estimation of the exposure level (*e.g.*, concentration) above which adverse effects can occur on organisms, and below which adverse effects are unlikely, *i.e.*, which can not be distinguished from background noise [OECD, 2006].

OECD (2006). *Current approaches in the statistical analysis of ecotoxicity data: a guidance to application.* Technical report ENV/JM/MONO(2006)18.

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Determination of a NOEC
Based on hypothesis tests

- ▶ **NOEC: "No Observed Effect Concentration"** :
  Maximal dose (or concentration) without any observed adverse effect.

- ▶ **LOEC: "Lowest Observed Effect Concentration"** :
  Minimal dose (or concentration) with an observed adverse effect.

Example for a continuous variable normally distributed:
Dunnett's test for a global comparison of the observed mean in every non-control groups to the observed mean in the control group
(with correction of α risk due to multiple comparisons).

What is ecotoxicology?
Dose-response modelling
Parameter inference

Introduction
Monospecific toxicity tests
Critical Effect Concentrations

Shortcomings of the NOEC

Severely criticized for multiple disadvantages

▶ necessarily one of the tested concentrations (hence strongly dependent on the experimental design);

▶ based on a wrong interpretation of the p-value (*absence of evidence is not evidence of absence*);

▶ strongly dependent on the sample size
$\rightarrow$ unprotective with small sample sizes: the lower the sample size, the higher the NOEC;

▶ cannot always be determined (e.g., if the first concentration leads to a significant difference);

▶ no uncertainty limits are associated.

sandrine.charles@univ-lyon1.fr   **BS-5-ECOSY**

What is ecotoxicology? Introduction
Dose-response modelling Monospecific toxicity tests
Parameter inference **Critical Effect Concentrations**

x% effective or lethal concentrations

Alternative to the NOEC, now **strongly recommended**.
$\rightarrow$ obtained by fitting a dose-response model to toxicity test
data at a chosen target time point, then deriving the dose which
corresponds to a given effect level (usually 10, 20 or 50%).
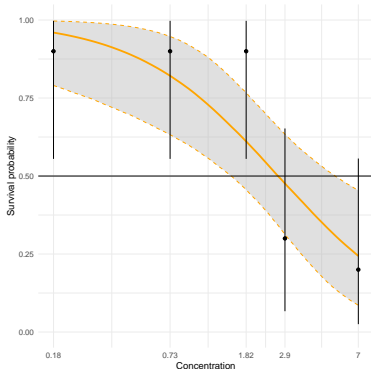
Advantages of EC_x or LC_x

▶ capture and account for
the whole dose-response
curve;

▶ slightly dependent on the
experimental design;

▶ may be associated to
uncertainty limits.

Shortcomings of EC_x or LC_x

▶ sometimes technical
difficulties when fitting;

▶ choice of a model;

▶ choice of an effect level x;

▶ choice of the exposure
duration.

What is ecotoxicology?   Introduction
Dose-response modelling   Monospecific toxicity tests
Parameter inference   **Critical Effect Concentrations**

Example of LC_x estimation

Use of survival data at the end of the experiment (day 21)

	median	2.5%	97.5%
LC5	0.22	0.0074	0.71
LC10	0.41	0.033	1.04
LC20	0.82	0.16	1.6
LC50	**2.67**	**1.50**	**5.3**

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

Table of content

What is ecotoxicology?

Dose-response modelling

Parameter inference

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Detailed content

Dose-response modelling
General points
Deterministic part
Stochastic part

What is ecotoxicology?	**General points**
Dose-response modelling	Deterministic part
Parameter inference	Stochastic part

Dose or Concentration-response or effect relationships?

A few vocabulary:

- ▶ **Dose** refers to the internal concentration, *i.e.*, the amount of toxicant **within** the body of organisms. But in ecotoxicology, only the **exposure concentration** is usually known.

  → We rather speak about **concentration**-response or effect relationships.

- ▶ Concentration-**response** relationships refer to the link between the exposure concentration and the proportion of individuals responding with an all-or-none effect.

- ▶ Concentration-**effect** relationships refer to the link between the exposure concentration and the magnitude of the induced biological change, measured in appropriate units.

What is ecotoxicology?	**General points**
Dose-response modelling	Deterministic part
Parameter inference	Stochastic part

Definition

A **concentration-response/effect relationship** is a simple X-Y graph relating increasing levels of exposure (X) to the response/effect (Y) **at a certain exposure time**.

Examples of **responses**:

▶ Quantal data, expressed as proportion or probability (*e.g.*, mortality or immobilization).

Examples of **effects**:

▶ Ordered descriptive categories (*e.g.*, severity of a lesion);
▶ Counts (*e.g.*, reproduction products like eggs or clutches);
▶ Continuous measurements (*e.g.*, body size).

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference Stochastic part

What is a regression model?

From concentration-response/effect experiments, if there is a reasonable number of concentrations (usually ≥ 5) of the toxicant and a reasonably well-behaved response/effect, it is straightforward to **fit a regression model**.

A regression model relating a dependent variable Y (the response or the effect) to an explanatory variable X (the concentration) is composed of two parts:

1. a **deterministic part**, which describes the mean value (or curve) (*e.g.*, a log-logistic model);
2. a **stochastic part**, which represents the distribution around the mean curve (*e.g.*, a normal distribution).
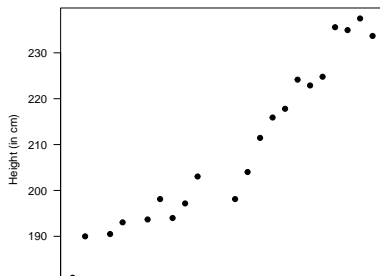
Nevertheless, each part depends on the nature of data to analyze.

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference Stochastic part

To make it simple

The Gaussian linear regression

Build a linear model between an **observed** quantitative continous variable Y and a **controled** quantitative continous variable X.

Example of the best high jump at the Olympic Games from 1896 to 1992.

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference Stochastic part

Best high jump at the Olympic Games
The model

$$Y_i = \alpha X_i + \beta + \varepsilon_i \quad \text{with} \quad \varepsilon_i \underset{i.d.}{\sim} \mathcal{N}\left(0, \sigma^2\right)$$

The linear function $f(x) = \alpha x + \beta$ is the deterministic part of the model, while ε_i stand for the stochastic part, assuming a Gaussian (or normal) distribution of the observations Y_i.

This model can equivalently be written as follows:

$$Y_i \sim \mathcal{N}\left(f(X_i), \sigma^2\right) \quad \text{with} \quad f(X_i) = \alpha X_i + \beta$$

In total, there are 3 parameters to estimate: α, β and σ.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Detailed content

Dose-response modelling
General points
Deterministic part
Stochastic part

What is ecotoxicology?	General points
Dose-response modelling	**Deterministic part**
Parameter inference	Stochastic part

Four shapes to describe dynamic in life science

From http://bioassay.dk/bioassay/

sandrine.charles@univ–lyon1.fr   **BS-5-ECOSY**

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
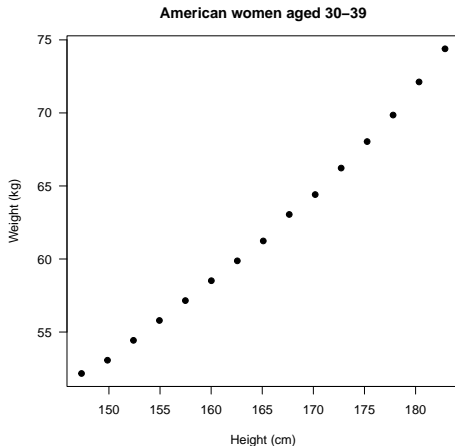**Parameter inference** **Stochastic part**
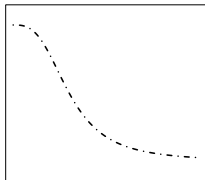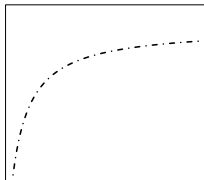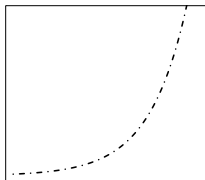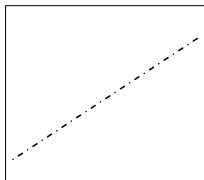
Example of **linear** relationship

Data on the relationship of average heights and weights for American women aged 3039.

$(n = 15)$

American women aged 30–39

Weight (kg) / Height (cm)

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
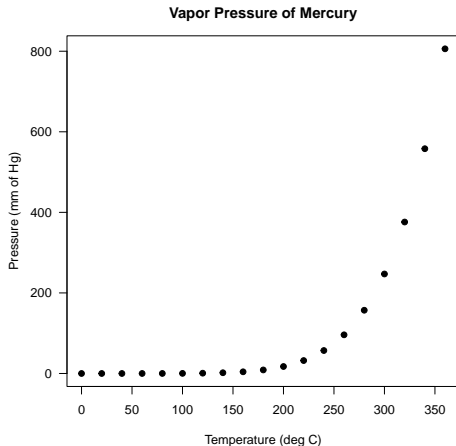**Parameter inference** **Stochastic part**

Example of **exponential** relationship

Data on the relation-
ship between temper-
ature in degrees Cel-
sius and vapour pres-
sure (in millimetres of
mercury).

$(n = 19)$

Vapor Pressure of Mercury

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

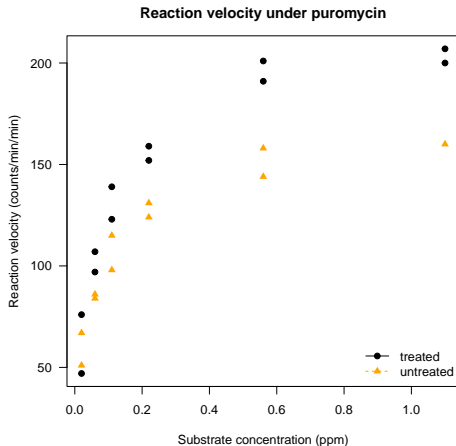## Example of **hyperbolic** relationship

Data on reaction velocity versus substrate concentration in an enzymatic reaction involving untreated cells or cells treated with puromycin (Michaelis-Menten).
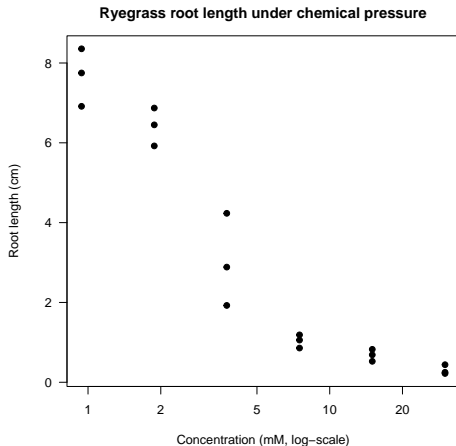
$(n = 12, n = 11)$

Reaction velocity under puromycin

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

Example of **sigmoidal** relationship

Data from a single dose-effect relationship between root lengths of perennial ryegrass (*Lolium perenne* L.) and concentration of ferulic acid.

($n = 24$)

Ryegrass root length under chemical pressure

Root length (cm)

Concentration (mM, log−scale)

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

Sigmoidal deterministic part

Among sigmoidal curves, the **log-logistic model** is the most commonly concentration-response/effect model used in (eco)toxicology [Ritz, 2010]:

$$Y = c + \frac{d - c}{1 + exp(b.\log(\frac{X}{e}))}$$
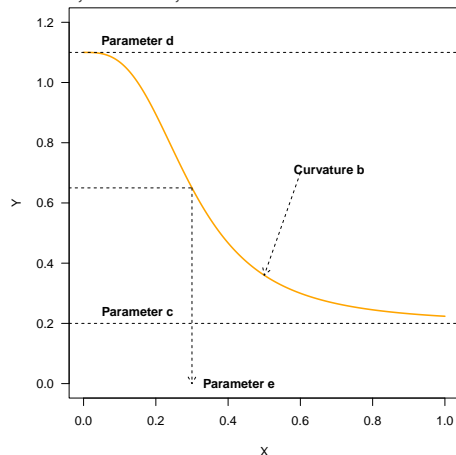
$$\Updownarrow$$

$$Y = c + \frac{d - c}{1 + (\frac{X}{e})^b}$$

b, c, d, e are positive, all with a geometric meaning.

Ritz C. (2010) Toward a unified approach to dose-response modeling in ecotoxicology. *Environmental Toxicology and Chemistry*, 29(1): 220229.

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

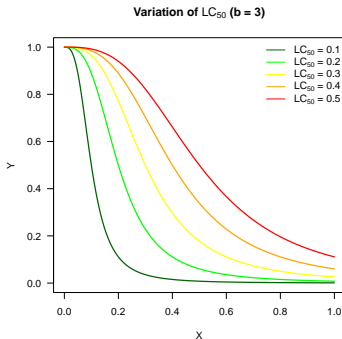# The log-logistic model - Graph

$b = 3, c = 0.2, d = 1.1$ and $e = 0.3$.
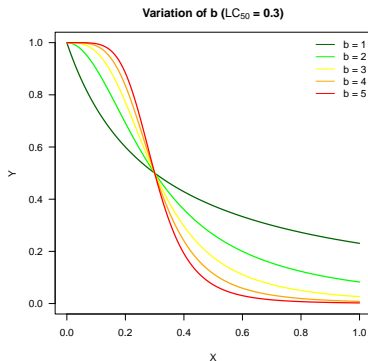

$e = 0.3$ (arbitrary unit)

In case of survival, d corresponds to the natural mortality (may be fixed to 1) and c is fixed to 0.

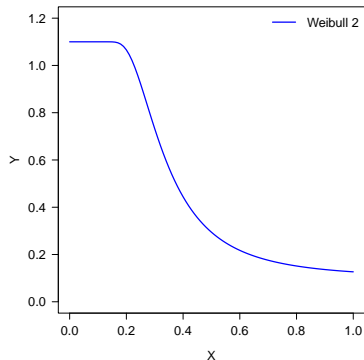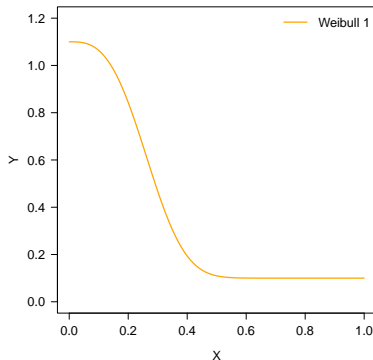|	
What is ecotoxicology?	**General points**
Dose-response modelling	**Deterministic part**
Parameter inference	*Stochastic part*

The log-logistic model - Morphology

Case of survival, with $c = 0$ and $d = 1$: $e = LC_{50}$.

What is ecotoxicology?   **General points**
Dose-response modelling   **Deterministic part**
Parameter inference   **Stochastic part**

Other sigmoidal models
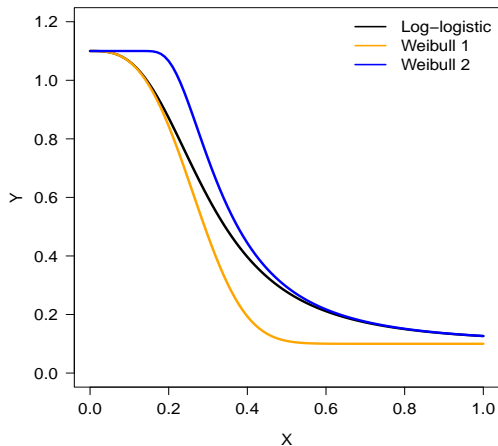
Many other models exist to describe sigmoidal shapes of dose-response curves as for example the Weibull's models:

$$Y = c + (d - c)e^{-(\frac{x}{e})^b} \quad \text{or} \quad Y = c + (d - c)(1 - e^{-(\frac{x}{e})^b})$$

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Comparison of sigmoidal curves

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Detailed content

Dose-response modelling
General points
Deterministic part
Stochastic part

sandrine.charles@univ-lyon1.fr **BS-5-ECOSY**

What is ecotoxicology?   General points
Dose-response modelling   Deterministic part
Parameter inference   **Stochastic part**

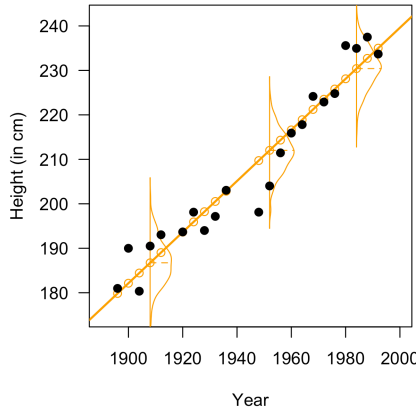# What is the stochastic part of a model?

The stochastic part

models the probability distribution **around** the average tendency of the data.

It depends on the nature of the data (namely quantal, discrete or continuous data)

What is ecotoxicology?	General points
Dose-response modelling	Deterministic part
Parameter inference	**Stochastic part**

What is the stochastic part of a model?

Example with the Gaussian linear regression

sandrine.charles@univ-lyon1.fr · BS-5-ECOSY

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Quantal data - description

Quantal (or binary) data arise when a particular property is recorded to be present or absent in each individual (*e.g.*, an individual shows an effect or it does not show an effect).

Therefore, these data can exhibit **only two states**.

Typically, quantal data are presented as **the number of individuals** showing the property (*e.g.*, mortality) out of a total number of individuals observed in each experimental unit.

Although this can be expressed as a fraction, **the total number of individuals cannot generally be omitted**.

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

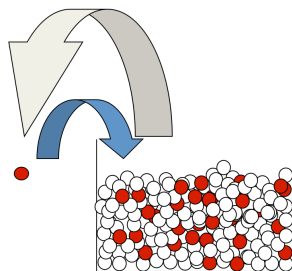## Quantal data - Binomial stochastic part

With **quantal** experimental data (*e.g.*, survival data), the
stochastic part of the model is necessarily **binomial**.

Observations are then described by a model of the following
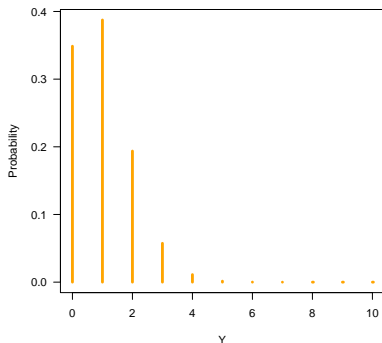form:

$$Y \sim \mathcal{B}\left(p(X, \theta), n\right)$$

where $p(X, \theta)$ is the **probability of success** (*e.g.*, survival
probability) as described by one of the dose-response models
(e.g., log-logistic or Pires-Fox), X is the exposure concentration,
θ is the parameter vector and n is the total number of
individuals.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Principle of a binomial process

- ▶ A **binomial** statistical experiment consists of n repeated trials, each trial resulting in just two possible outcomes: success (of probability p) or failure.

- ▶ Trials are independant.

- ▶ Let Y be the number of successes resulting from a binomial experiment, then $Y \sim \mathcal{B}(p, n)$.

- ▶ The probability distribution of Y is a **binomial distribution** of mean $p \times n$ and variance $p \times (1 - p) \times n$.

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

Binomial probability distribution - Graph

$p = 0.1$ and $n = 10$

▶ $P(Y = 1)$ is the highest

▶ $P(Y = 8)$ and $P(Y = 9)$ are very low

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

Binomial probability distribution - Morphology

$n = 10$

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**
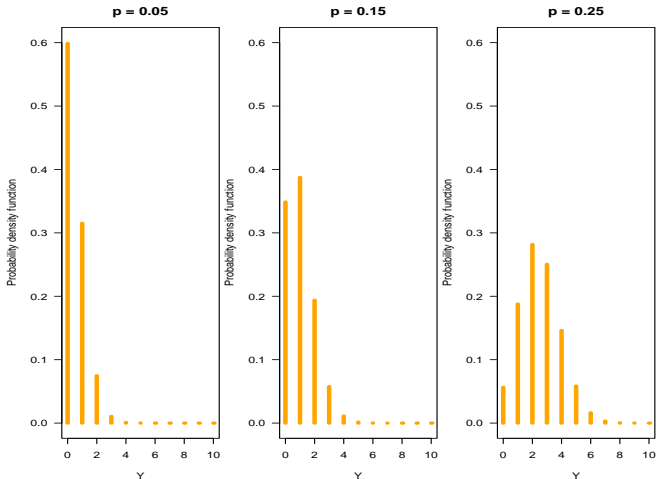
Discrete data - description

Data are **discrete** if there are only a finite number of possible values or if there is a space on the number line between each two possible values.

Discrete data are usually obtained when we are **counting something** (using whole numbers).

They are also called **count data**.

Typical discrete data are **reproduction data**.

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

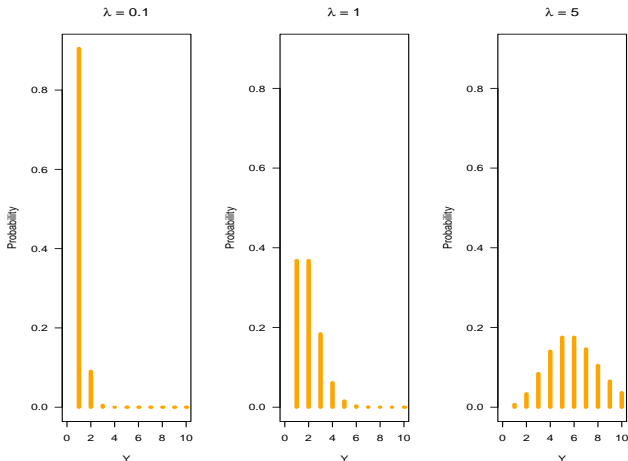## Discrete data - Poisson stochastic part

Count data are usually modelled using a **Poisson distribution**. If N is the number of reproduction outputs (*e.g.*, eggs or clutches) at concentration X, then:

$$N \sim \mathcal{P}(\lambda)$$

with $\lambda = f(X, \theta)$ the mean of the Poisson distribution, that is the predicted mean value from the log-logistic model, the Pires-Fox model or any other deterministic part: X is the exposure concentration and θ the parameter vector.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Poisson probability distribution - Vizualization

λ denotes both the mean and the variance of the distribution.

sandrine.charles@univ-lyon1.fr   **BS-5-ECOSY**

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

Continuous data - description

Data are **continuous** when they can (theoretically) take any value in an open interval.

Examples include measurements of length, body weight, etc.

Due to practical reasons the measured resolution depends on the quality of the measurement device.

Typically, **continuous data have a dimension** (*e.g.*, g, cm, g.L^{-1}).

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

Continuous data - Gaussian stochastic part

With **continuous** experimental data, the currently encountered stochastic part is **normal** (or **Gaussian**), even if other stochastic parts may be sometimes more appropriate.

Observations are then described by a model of the following form:
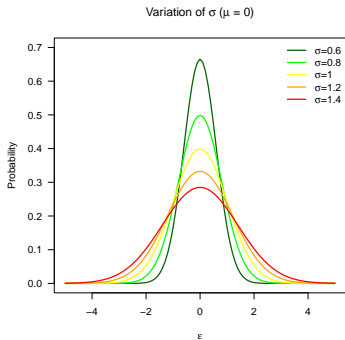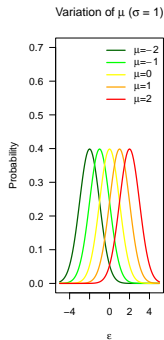
$$Y \sim \mathcal{N}(f(X, \theta), \sigma)$$

with X the independant variables (*e.g.*, the concentration), $f(X, \theta)$ the log-logistic or the Pires-Fox model and θ the vector of model parameters.
This expression can equivalently be written as follows:

$$Y = f(X, \theta) + \varepsilon \quad \text{with} \quad \varepsilon \sim \mathcal{N}(0, \sigma)$$

The normal distribution is a **bell-shaped** probability density function with two parameters: mean μ and variance σ^2.

What is ecotoxicology?

Dose-response modelling

Parameter inference

General points

Deterministic part

Stochastic part

The bell-shaped Gaussian probability distribution

What is ecotoxicology? General points
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

In brief: a wide variety of models

1. A **deterministic part**: linear or non-linear,
   and its associated parameters:
   for example, (α, β) or (b, c, d, e);

2. A **stochastic part**: Gaussian or not,
   and its associated parameters:
   for example, p (binomial), λ (Poisson) or σ (normal).

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

A battery of regression types

▶ **Gaussian linear regression**: simple linear regression, polynomial regression, multiple regression,...;

▶ **Generalized linear regression**: logistic regression, Poisson regression, multinomial logit regression, probit regression,...;

▶ **Gaussian non-linear regression**: least-square regression, simple or multiple;

▶ **Generalized non-linear regression**.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
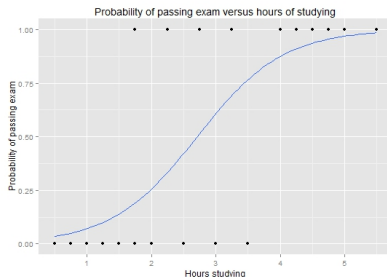**Stochastic part**

Logistic regression

The logistic regression deals with **quantal data**. It is used to model the probability of "success" p via a relationship between the **logit** (or the log-odd) of p and a **linear combination** of one or more independent variables ("predictors"):

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_0 X_1 + \beta_2 X_2 \ldots + \beta_n X_n$$

The quantity $\frac{p}{1-p}$ is called an **odd ratio**: the ratio of the probability of success over the probability of failure.

Remark: *logit* stands for ***log***istic un***it***.

What is ecotoxicology?	General points
Dose-response modelling	Deterministic part
Parameter inference	Stochastic part

Logistic regression: example

Probability of passing exam versus hours of studying

$logit$(Passing exam) $= 1.5046$ Hours $- 4.0777$ with $p - value = 0.0167$

According to the model, 2 hours of study leads to an estimated probability of passing the exam of 0.26, while the probability becomes 0.87 with 4 hours of study.

Source: https://en.wikipedia.org/wiki/Logistic_regression

What is ecotoxicology? **General points**
Dose-response modelling **Deterministic part**
Parameter inference **Stochastic part**

Poisson regression

The Poisson regression deals with **count data**. The response
variable Y is assumed to follow a Poisson distribution of
parameter λ, and the logarithm of λ is modeled by a **linear
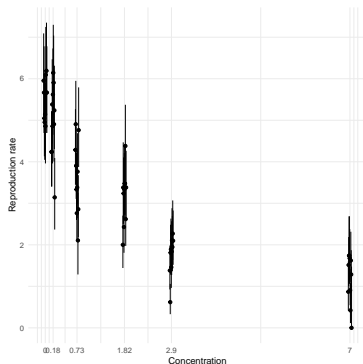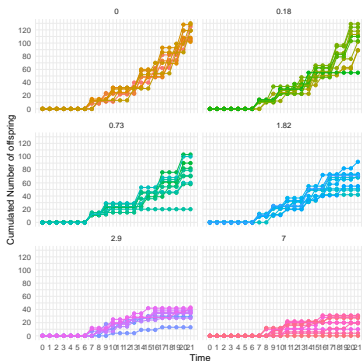combination** of one or more independent variables
("predictors"):

$$\ln(\lambda) = \beta_0 + \beta_0 X_1 + \beta_2 X_2 \ldots + \beta_n X_n$$

Under this model, the mean is assumed to be equal to the
variance.

Remark: A Poisson regression model is sometimes called a
log-linear model.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Deterministic part
Stochastic part

Poisson regression

Example with a nonlinear $X - Y$ relationship

sandrine.charles@univ-lyon1.fr   **BS-5-ECOSY**

What is ecotoxicology? **General points**
Dose-response modelling Deterministic part
Parameter inference **Stochastic part**

Regression analysis: for what?

▶ Widely used for prediction;

▶ To understand which ones of the independent variables are related to the dependent variable;

▶ To infer causal relationships between the independent and dependent variables; but **caution is advisable**:
for example, correlation does not imply causation.

Regression analysis: how to do in practice?

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Table of content

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Detailed content

Parameter inference
  General points
  Frequentist vs Bayesian framework

sandrine.charles@univ-lyon1.fr **BS-5-ECOSY**

What is ecotoxicology?
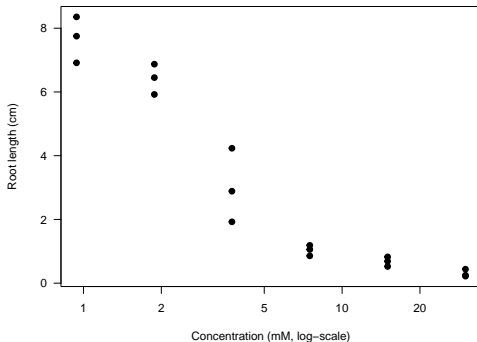Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

What is inference ?

Inference usually implies to fit a model to observed data.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

What is inference ≡ Get parameter estimates

Several criteria may provide the best fit parameter values.

Concentration (mM, log–scale)
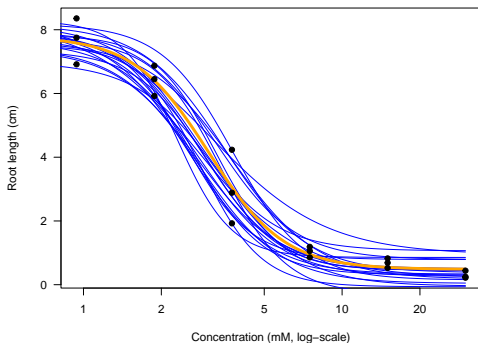
⇒ Search for best fit based on the concept of probability.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

What is a probability ?
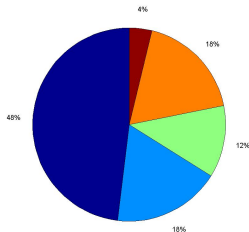
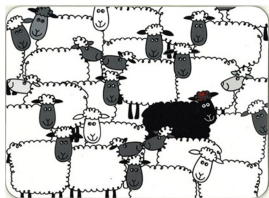

One word,
at least two definitions.

From a lecture by Marie Laure Delignette-Muller (http://www2.vetagro-sup.fr/ens/biostat/introBayesPreditox_juil14.pdf)

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Frequentist view of probability

In a frequentist perspective, the probability of an event is defined as the fraction of times that the event occurs in a very large number of trials.

Probability of being a black sheep ?

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
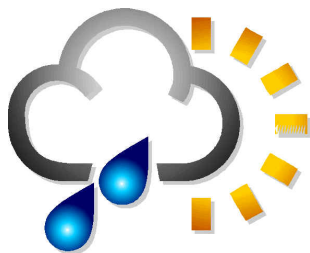Frequentist vs Bayesian framework

Bayesian view of probability

In a Bayesian perspective, the probability is seen as a degree of belief, a measure of uncertainty.

Probability of rain tomorrow ?

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Generalization to population

Inference also implies **generalization** from a sample to population, and the calculation of **uncertainty** in the estimated parameters, especially uncertainty due to sampling error.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework
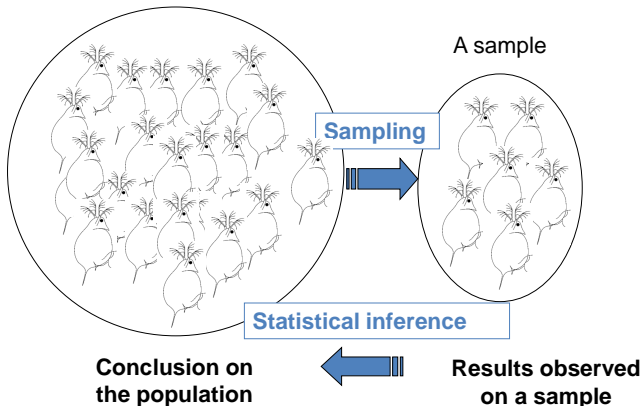
Let's take a very simple example

Estimation of a probability to survive for animals studied under fixed conditions
Data : $Y = 24$ survivals among $n = 100$ organisms

The model:

▶ No deterministic part
▶ Binomial stochastic part : $Y \sim \mathcal{B}(p, n)$

This model is characterized by **only one parameter**: p.
For the sake of generality, the vector of model parameters will be denoted θ hereafter.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Detailed content

Parameter inference

General points

Frequentist vs Bayesian framework

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Point estimate of θ: $\widehat{\theta}$

Frequentist framework

Parameter θ is assumed fixed but unknown

Parameter θ is estimated by one of the following methods:

▶ Maximum likelihood: $\max\limits_{\theta} P\left(Y \mid \theta\right)$;

▶ Moment matching;

▶ Minimization of sum of squared deviations.

In some cases, these different methods may lead to the same estimation of θ.

Example: the estimated survival probability is $\widehat{p} = \frac{24}{100} = 0.24$.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Interval estimate of θ: confidence interval

**The calculation of a confidence interval (generally a 95%
interval) is based on repeated sampling from the model :**

Definition

If we repeatedly obtain samples of size n from the population
and build a 95% confidence interval for each, we can expect
95% of the intervals to contain the true value of the parameter.

In average, among the 95% confidence intervals we obtained, **1
out of 20** does not contain the true value of the parameter to
estimate.

Example: the 95% confidence interval of the survival
probability is $\widehat{p} \pm 2\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$, that is $[0.15; 0.33]$.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

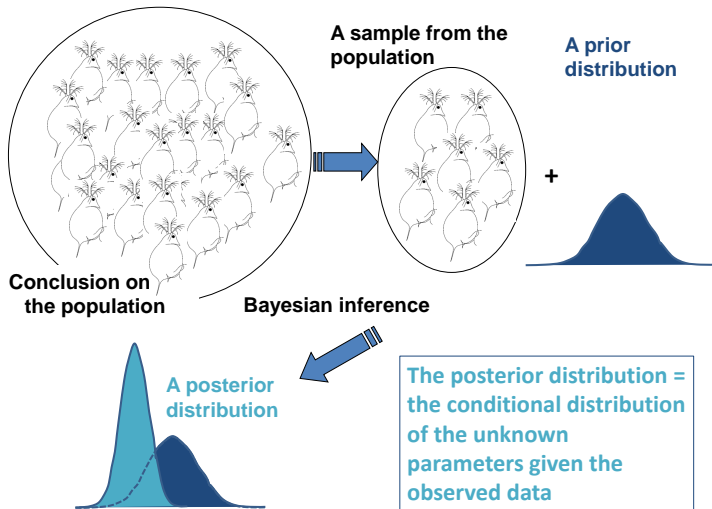# Bayesian estimation of θ

Bayesian framework

Parameter θ is supposed uncertain, and its uncertainty is characterized by a probability distribution (subjective meaning of a probability, degree of belief)

- Prior distribution: $P(\theta)$ more or less informative;
- Posterior distribution: $P(\theta|Y)$ calculated using the Bayes theorem, from the prior distribution and the likelihood function $P(Y|\theta)$;

$$P(\theta|Y) = \frac{P(Y|\theta) \times P(\theta)}{P(Y)} \propto P(Y|\theta) \times P(\theta)$$

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Principle of Bayesian inference

A sample from the population

A prior distribution

+

Conclusion on the population

Bayesian inference

A posterior distribution

The posterior distribution = the conditional distribution of the unknown parameters given the observed data

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Use of the posterior distribution for parameter estimation

▶ **Point estimate:**
Mean, median or mode of the posterior distribution

▶ **Interval estimate:**
Definition of a **credible interval** (or Bayesian confidence interval) from posterior distribution quantiles:
$\rightarrow 2.5\%$ and 97.5% quantiles for a 95% credible interval.
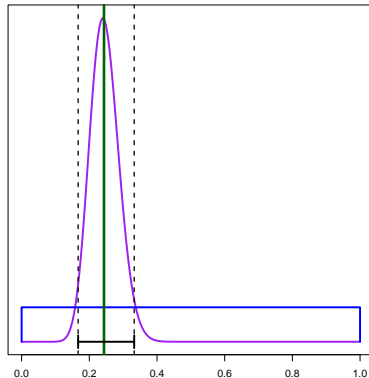**Easy interpretation**: the probability that the parameter lies in a 95% credible interval is 95%.

▶ Hypothesis tests:
It is **no more necessary** to calculate any p-value: one can make decisions directly from posterior distributions.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Example : Bayesian estimation of a survival probability

- ▶ Likelihood function: $\mathcal{B}(p, n)$
- ▶ Data: $Y = 24$ survivals out of $n = 100$
- ▶ Prior distribution : $U(0, 1)$ non informative
- ▶ Posterior distribution: analytically known in simple cases

- ▶ Point estimate : 0.24
- ▶ 95% credible

sandrine.charles@univ–lyon1.fr   BS-5-ECOSY

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Frequentist framework

- ▶ **Parameter θ is supposed fixed but unknown**;

- ▶ Parameter inference only uses observed data;

- ▶ Confidence intervals are based on repeated sampling from the model, the probability being associated to the relative occurrence frequency of an outcome.

Bayesian framework

- ▶ **Parameter θ is considered as a random variable, associated to a probability distribution**;

- ▶ Parameter inference uses both observed data and prior information (prior distribution);

- ▶ Credible intervals are defined from the posterior distribution and can be easily interpreted: 95% is the probability that the true parameter value lies within its 95% credible interval.

What is ecotoxicology?
Dose-response modelling
Parameter inference

General points
Frequentist vs Bayesian framework

Frequentist framework

- ▶ **Parameter θ is supposed fixed but unknown**;

- ▶ Parameter inference only uses observed data;

- ▶ Confidence intervals are based on repeated sampling from the model, the probability being associated to the relative occurrence frequency of an outcome.

Bayesian framework

- ▶ **Parameter θ is considered as a random variable, associated to a probability distribution**;

- ▶ Parameter inference uses both observed data and prior information (prior distribution);

- ▶ Credible intervals are defined from the posterior distribution and can be easily interpreted: 95% is the probability that the true parameter value lies within its 95% credible interval.