

Population Concepts and Clinical Trial Simulation

Robert R. Bies
Indiana University

- Population Analysis Concepts
- Clinical Trial Simulation
 - Components of a Clinical Trial Simulation
 - Public Health Question Example

Population Concepts

- What are we trying to describe?
 - Group characteristics?
 - Individual characteristics?
 - Explanatory variables within a group?

How do we describe groups?

- Central Tendency

- Mean

- The arithmetic average of all the available values of the descriptor (random effect)

- Median

- The descriptor value such that 50% of the remaining values are larger and 50% are smaller

- Mode

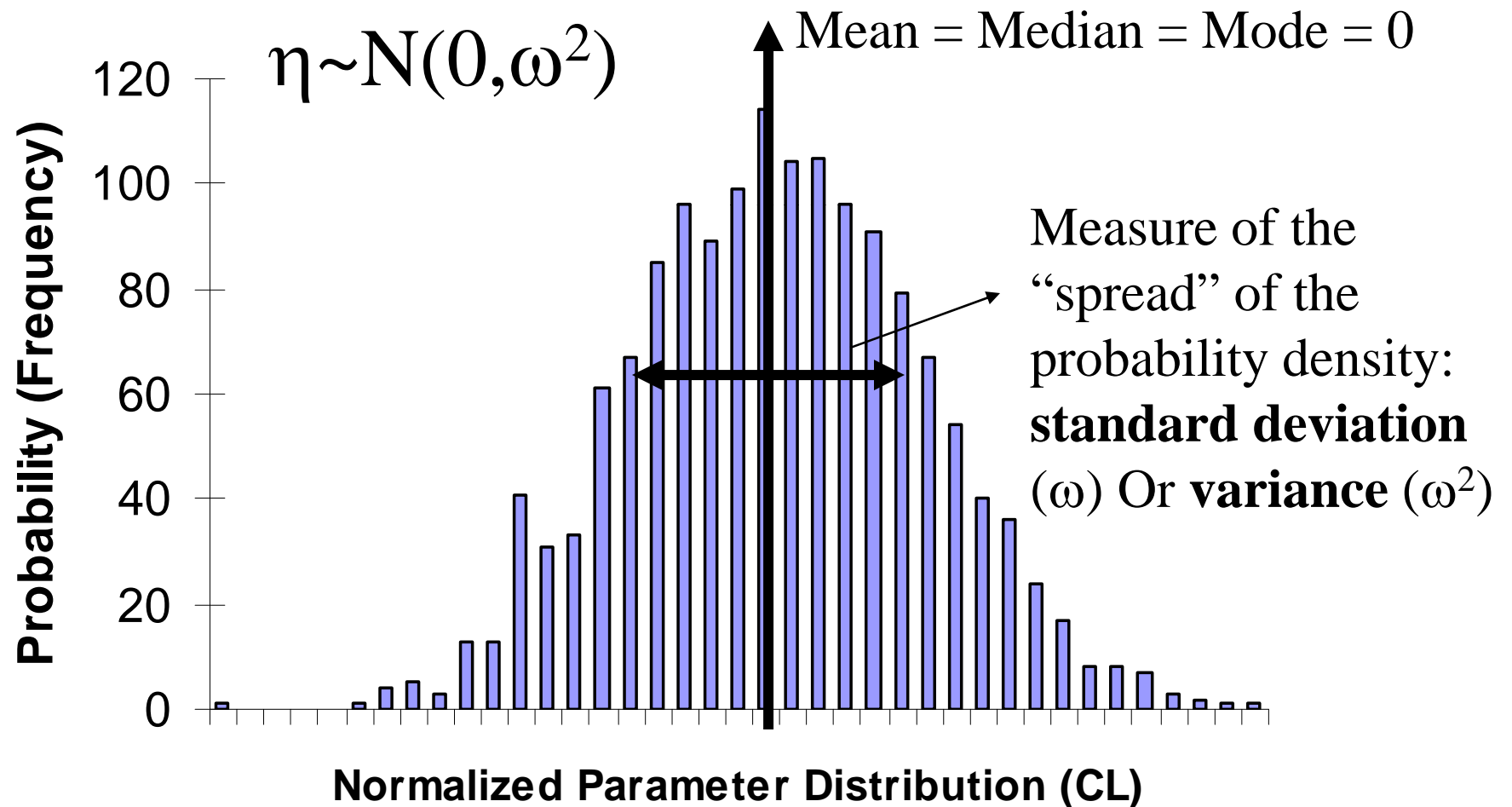
- The descriptor value at which the probability density attains its maximum value

How do we describe groups?

- Variability in the group
 - Standard Deviation
 - The average of the squared deviations of the random effects from the mean
 - Variance
 - The square root of the variance
 - Coefficient of Variation
 - The ratio between standard deviation and mean

A Distribution of Clearance Values

- Usually assumed Normal, centered at 0:



(also known as the infamous “bell curve”)

Approaches to Population Modeling

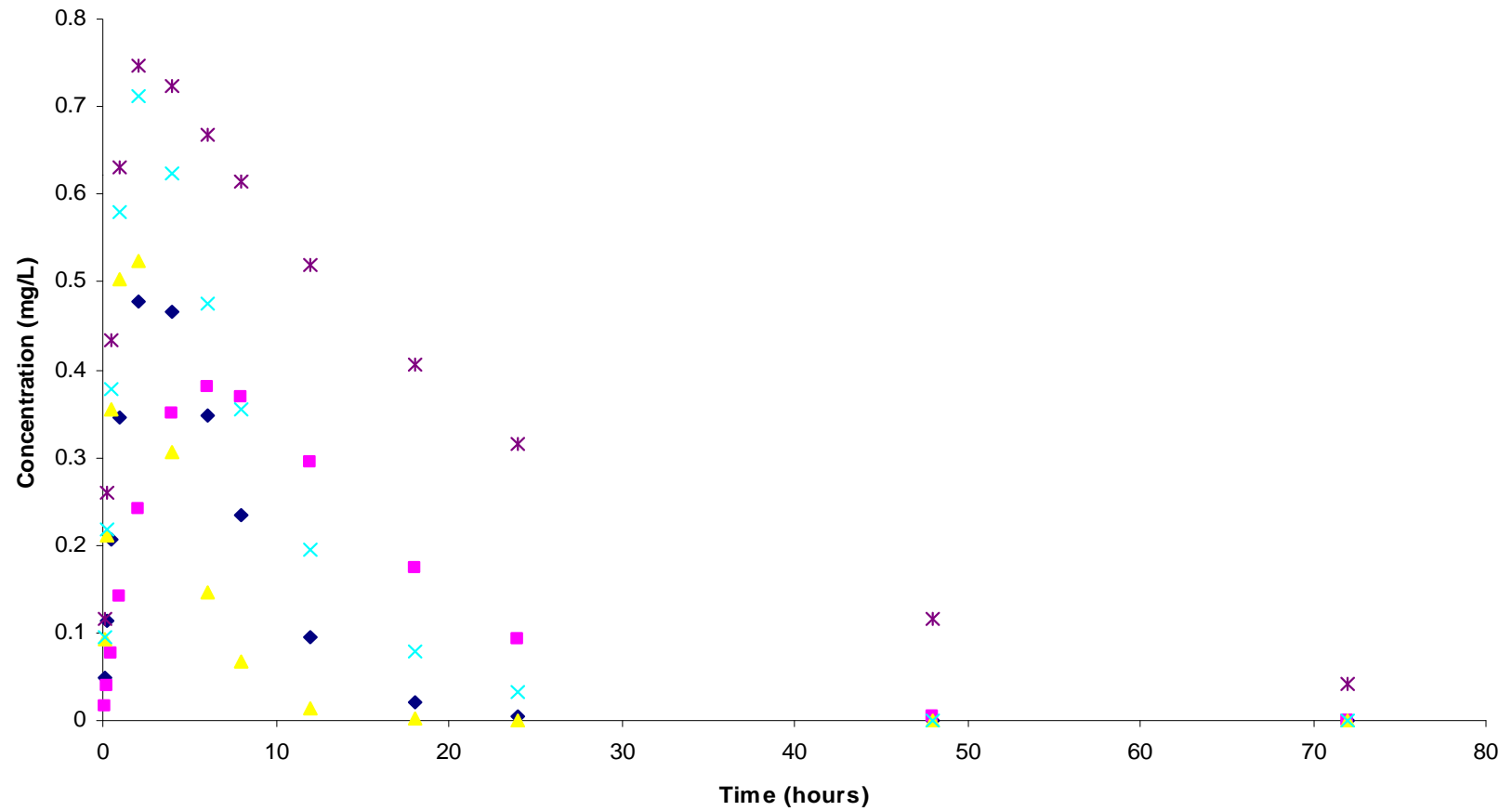
- Naïve Pooling
- Fitting Average Profile
- Standard Two-Stage
- Iterative Two-Stage
- Nonlinear Mixed Effects Modeling
- Bayesian Inference

Approaches – Naïve Pooling

- All data points are assumed to arise from a single individual
- A single function is fit to the combination of all individuals
- Ignores intra-individual variability as well as correlation in time within individuals

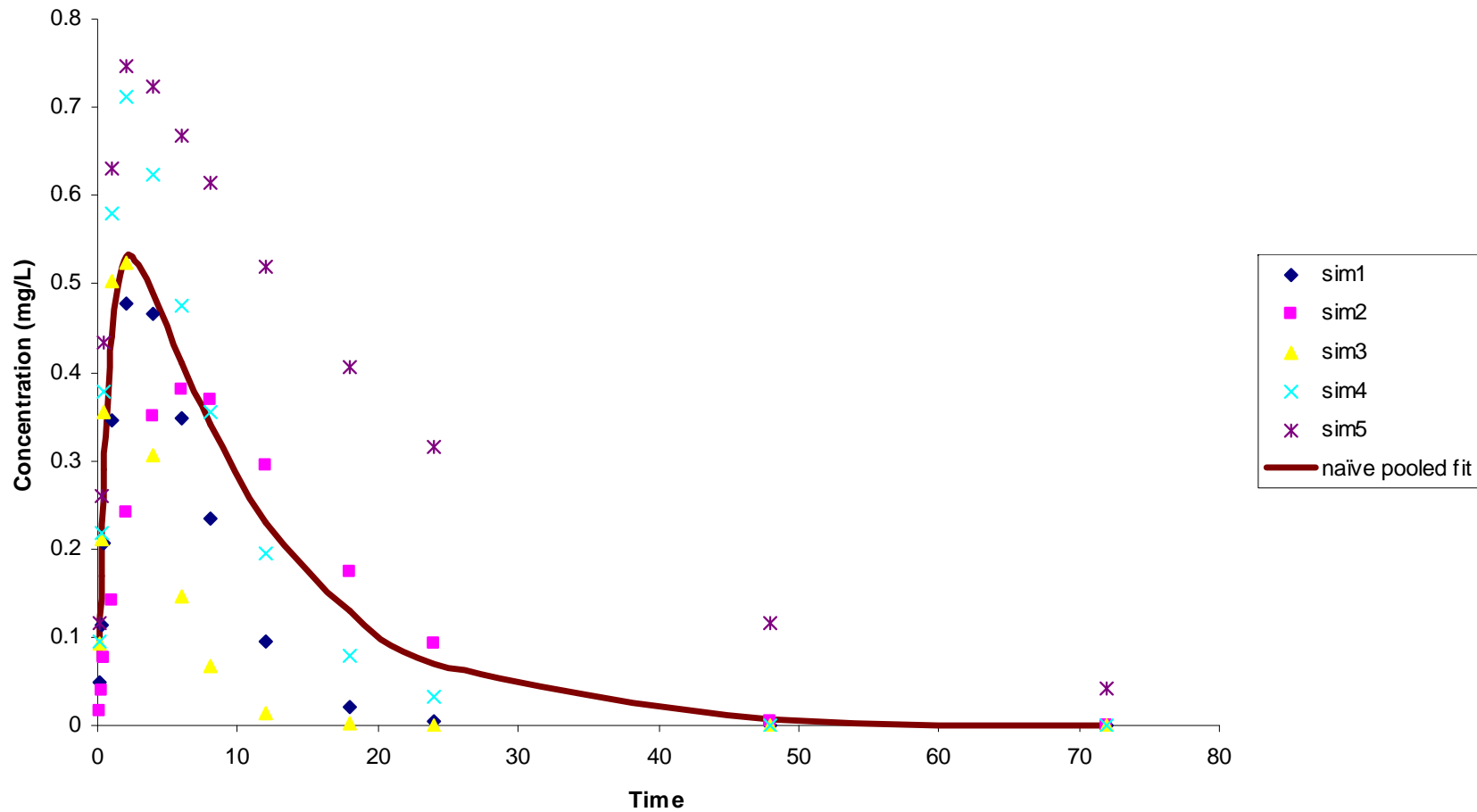
Naïve Pooling

Raw Data for PK profiles



Naïve Pooling

Naive Pooled Graph

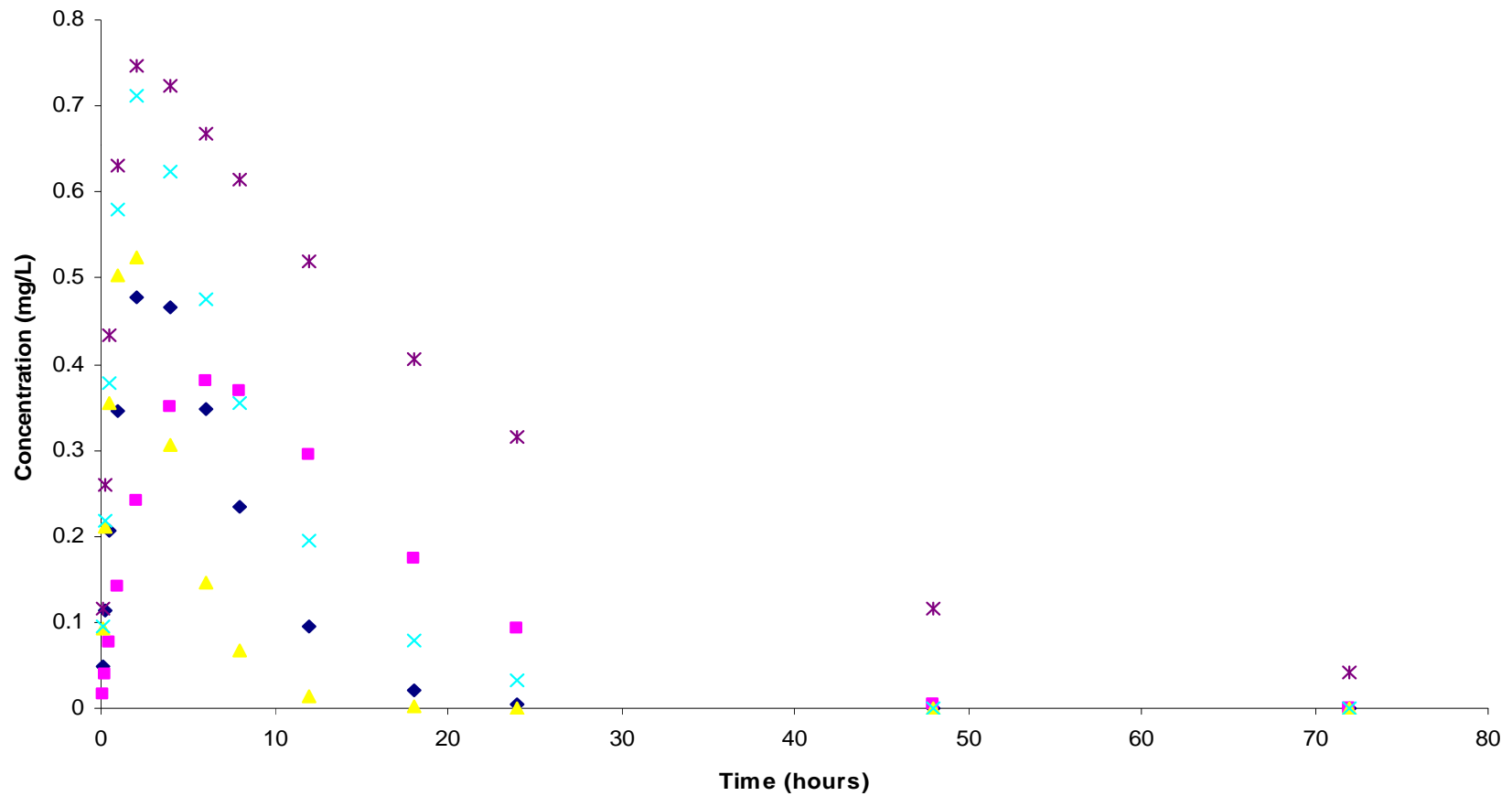


Approaches – Fitting the Average Profile

- Data points are averaged at each point in the measurement sequence
- Measurements must be made at the same times across individuals
- The averaged points across time are then used to fit or produce estimates for the pharmacokinetic model

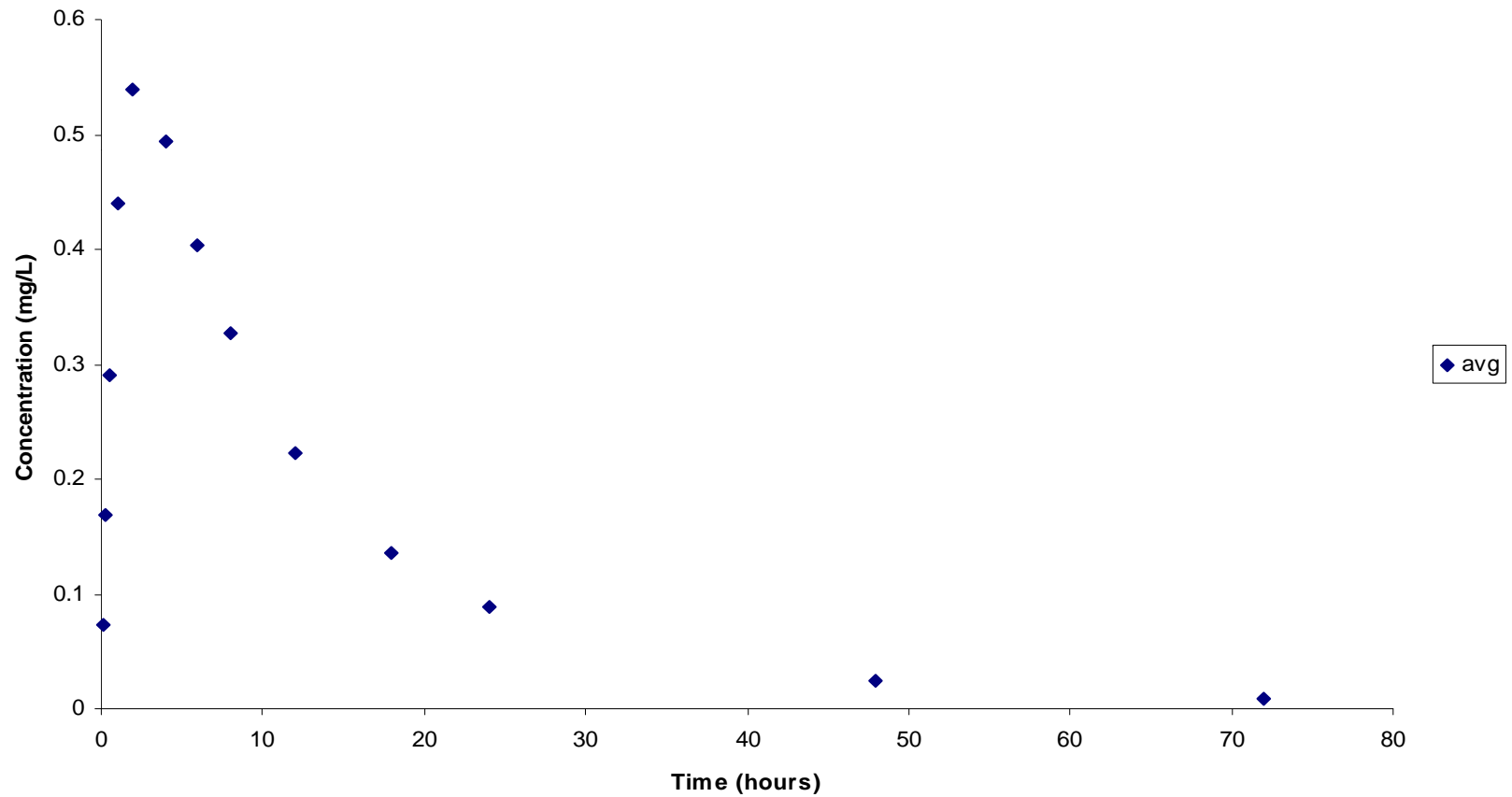
Fitting the Average Curve

Raw Data for PK profiles



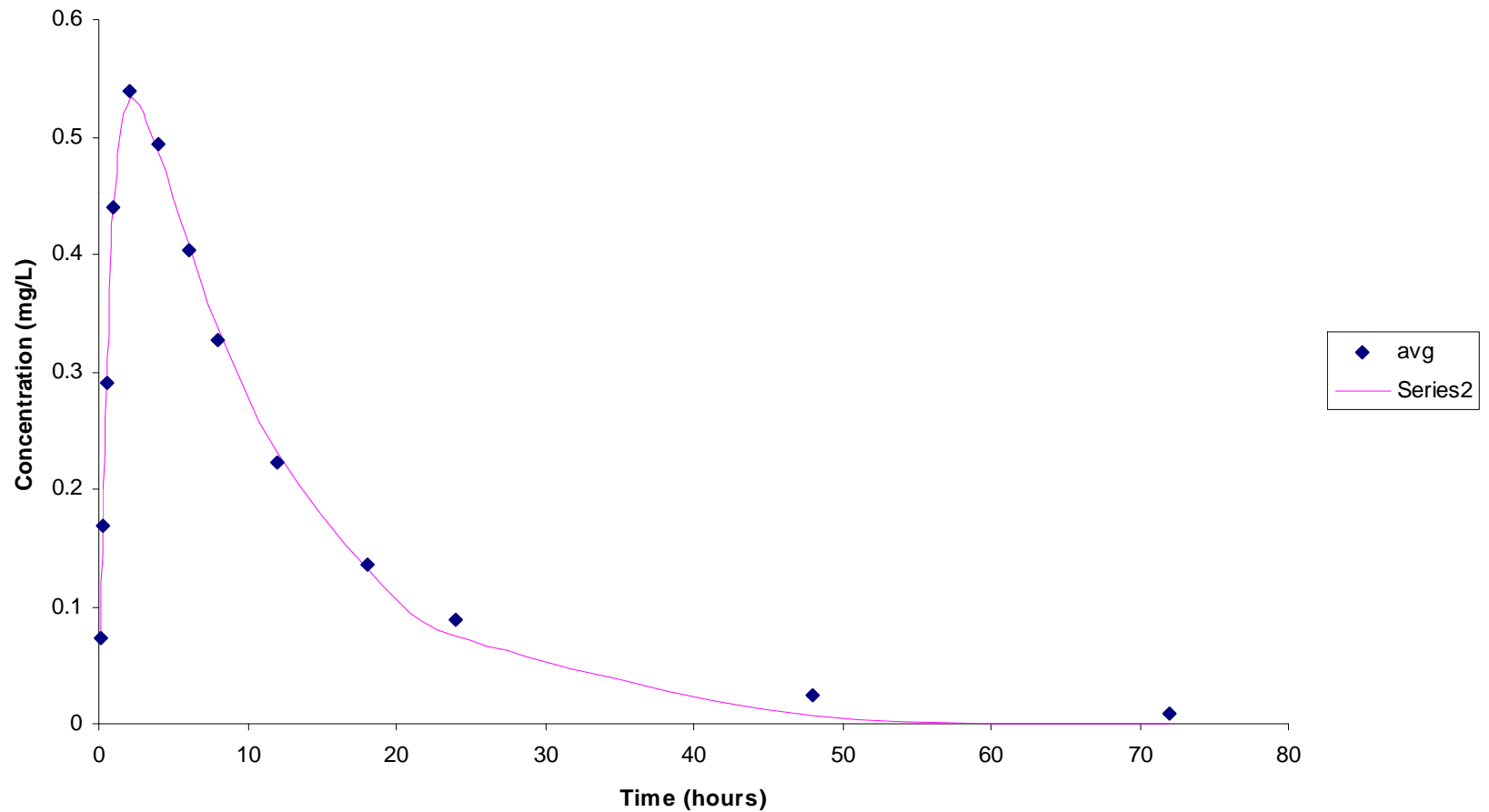
Fitting the Average Curve

Average Data Profile



Fitting the Average Curve

Average Data with Fit



PK Parameter Values

Naïve Pooled, Average and True

Summary Parameters

	True Average	Naïve Pooled	Average Fitted
CL	20	14.25	14.22
V	108	150.4	150.6
Ka	0.84	1.2	1.21

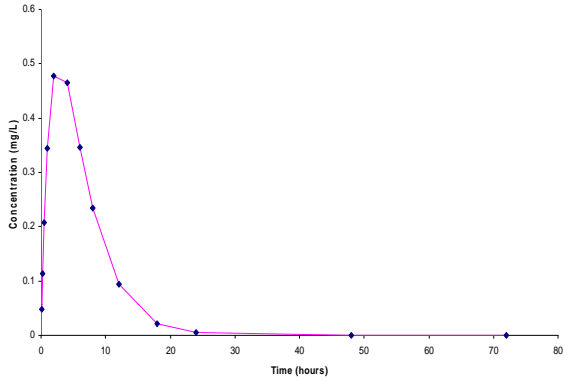
Approaches-Standard Two Stage

– Standard Two Stage

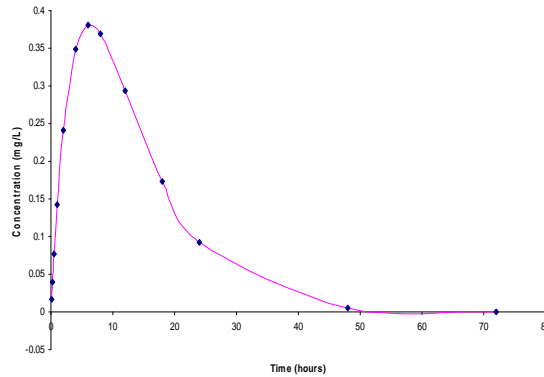
- Given complete PK profiles for a certain number of subjects
- each individual's PK parameters are estimated separately
- the sample mean and covariance of all the parameters is computed

Standard Two-Stage – Individual Data and Fit

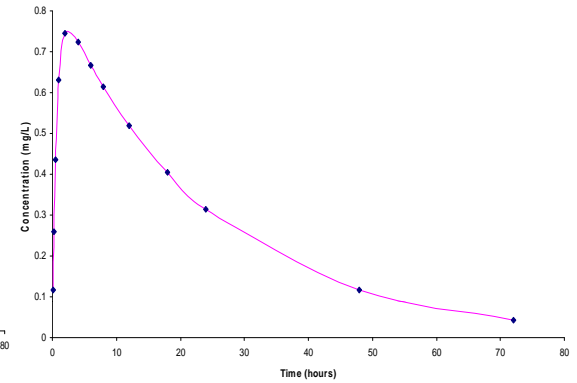
ID1 DATA AND FIT



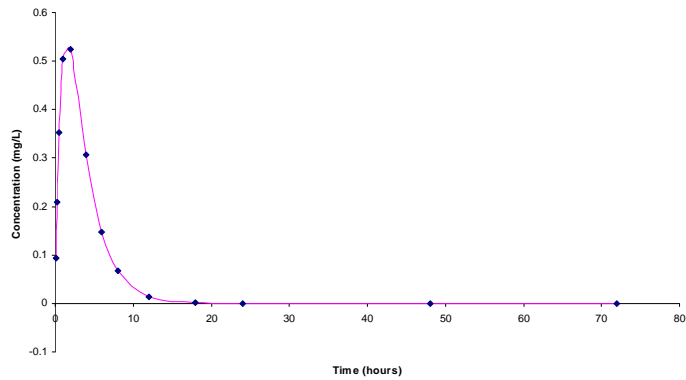
ID2



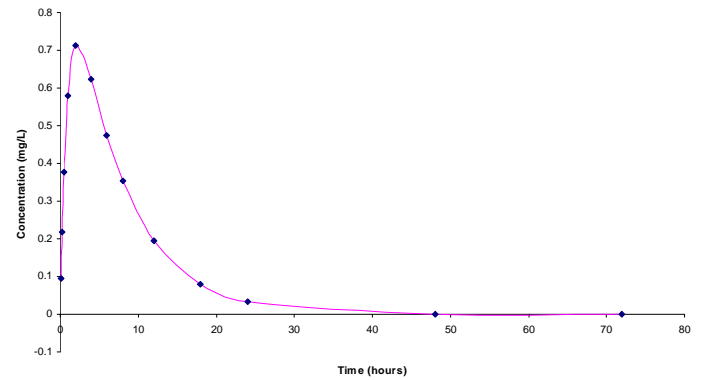
ID5



ID3

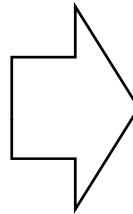


ID4



Standard Two-Stage

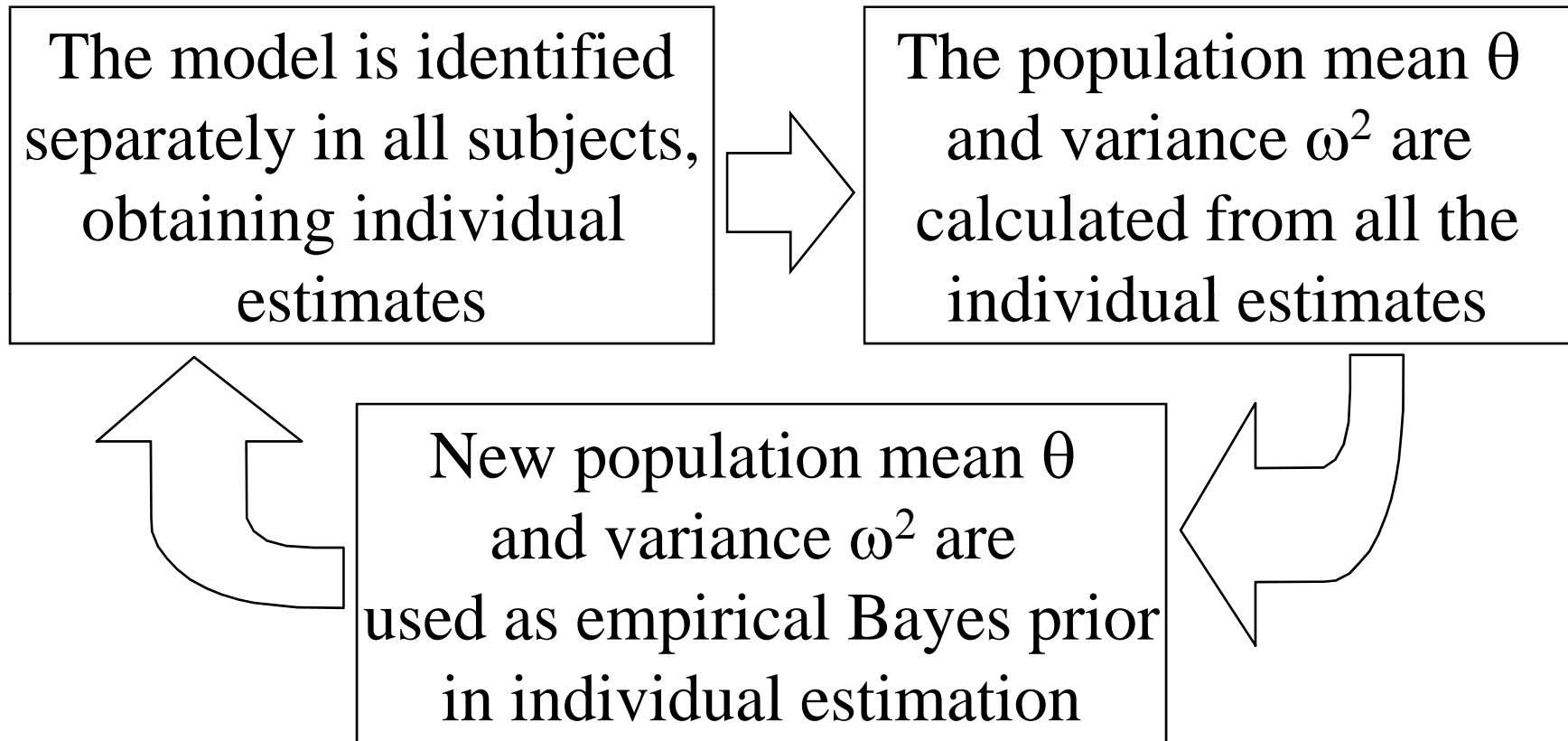
The model is identified separately in all subjects (e.g. with least squares), obtaining individual estimates of the parameters



The population mean θ and variance ω^2 are calculated as the sample mean and variance

- * Requires richly sampled data
- * Ignores the precision of the individual estimates
- * Overestimates the population variance

Iterative Two-Stage



- * It also works in some sparse data situations
- * Gives more reliable individual estimates

Parameter Values

Standard Two Stage Averages

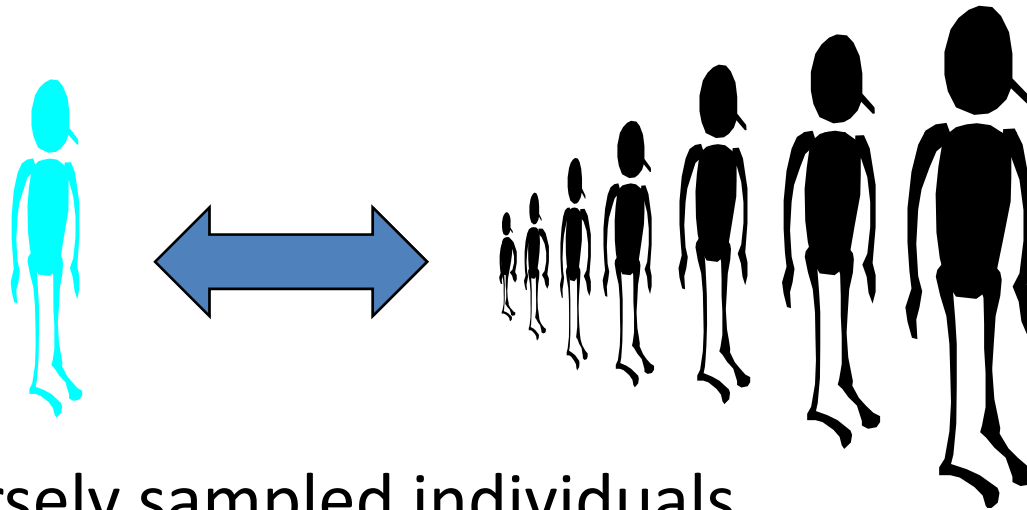
Summary
Parameters

	True Avg	STS Avg	Stdev True	Stdev STS	Naïve Pooled	Average Fitted
CL	20	20	13.22	13.23	14.25	14.22
V	108	108	10.95	10.94	150.4	150.6
Ka	0.84	0.84	0.5	0.5	1.2	1.21

From Individual to Population: Statistical Modeling

An Individual

A Population



- Many sparsely sampled individuals
⇒ *population* rather than *individual* response
- Statistical issues become paramount
- Goal: characterize and model variability sources

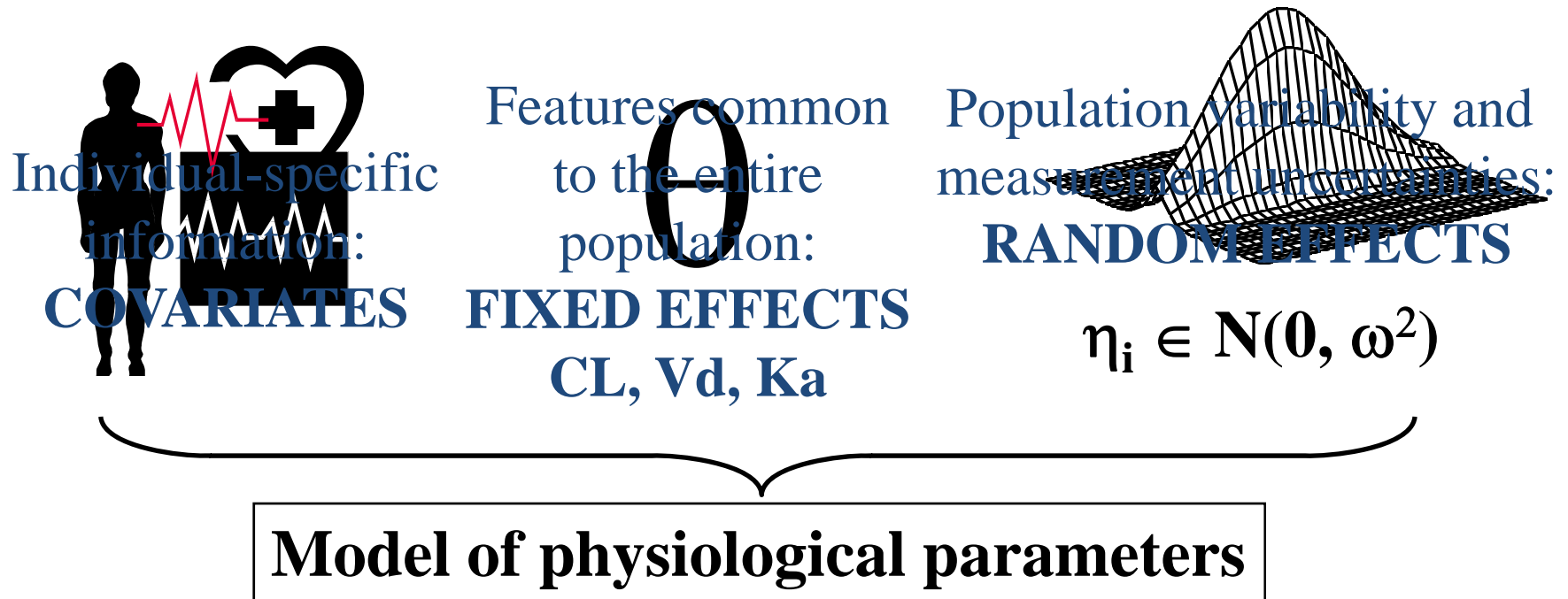
Stochastic Model

- This model provides answers to the following:
 - What is the extent of variability of PK parameters between subjects? [*between-subject variability*]
 - What is the extent of variability of model parameters in the same subject studied on multiple occasions? [*between-occasion variability*]
 - What is the extent of model misspecification and of unexplained variability in the concentration or effect measurements? [*residual unknown variability*]

Approaches – Nonlinear Mixed Effects Modeling

- Comprises Fixed and Random Effects
- Determine parameters from sparse data
- Explicitly determine inter-individual and intra-individual variability

Fixed and Random Effects



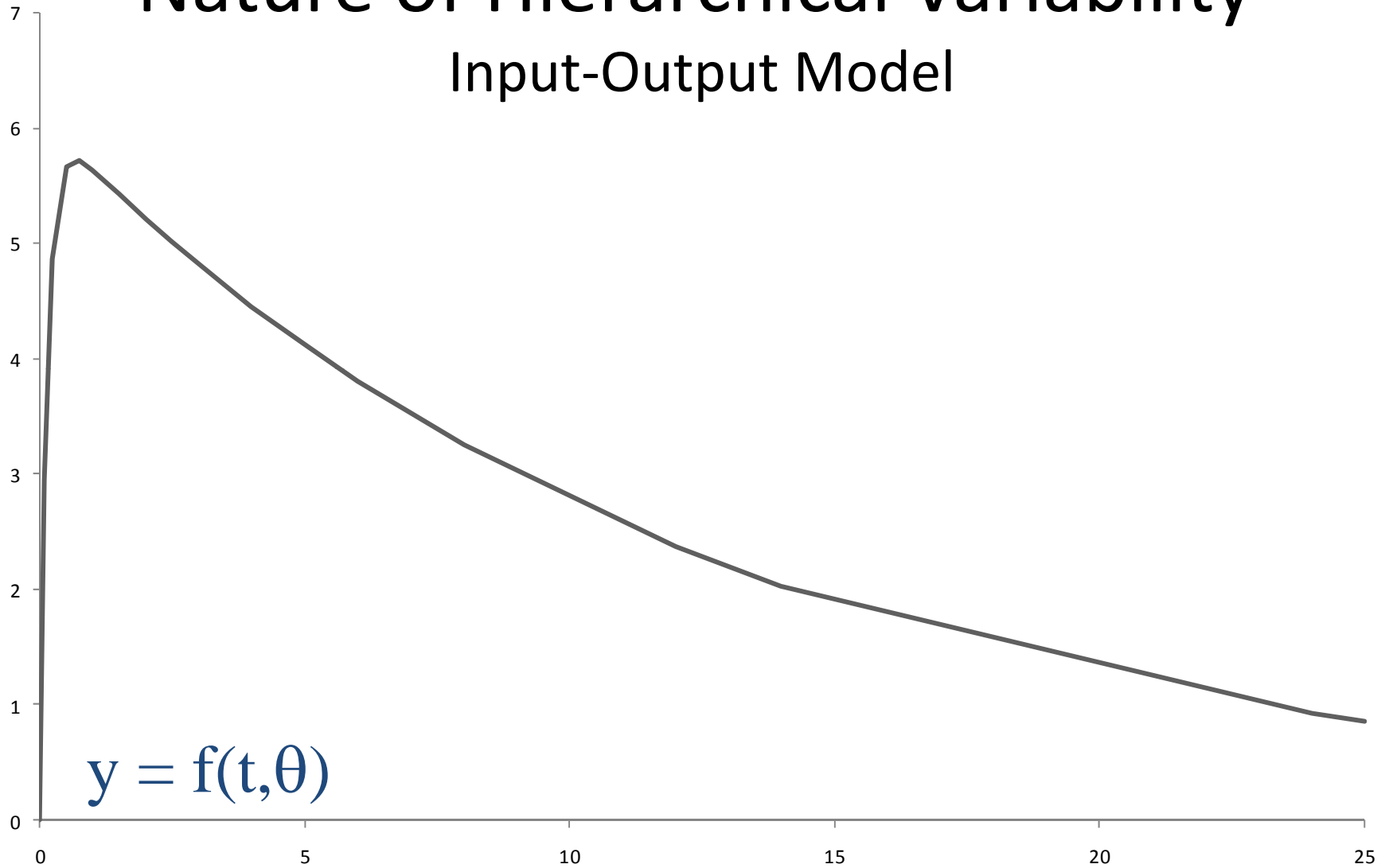
- This is often denoted as a *mixed effect model* (mixed = random + fixed, or varying + constant)
- It is a combination of the I/O and the stochastic models

Overview of Stochastic Models

- Types of Stochasticity/Variability
 - **Between-Subject Variability = BSV**
 - inter-subject variability
 - between-individual variability
 - **Between-Occasion Variability = BOV**
 - inter-occasion variability
 - **Residual Unknown Variability = RUV**
 - intra-subject variability
 - intra-individual variability
 - within-subject variability

Nature of Hierarchical Variability

Input-Output Model



Between-Subject Variability

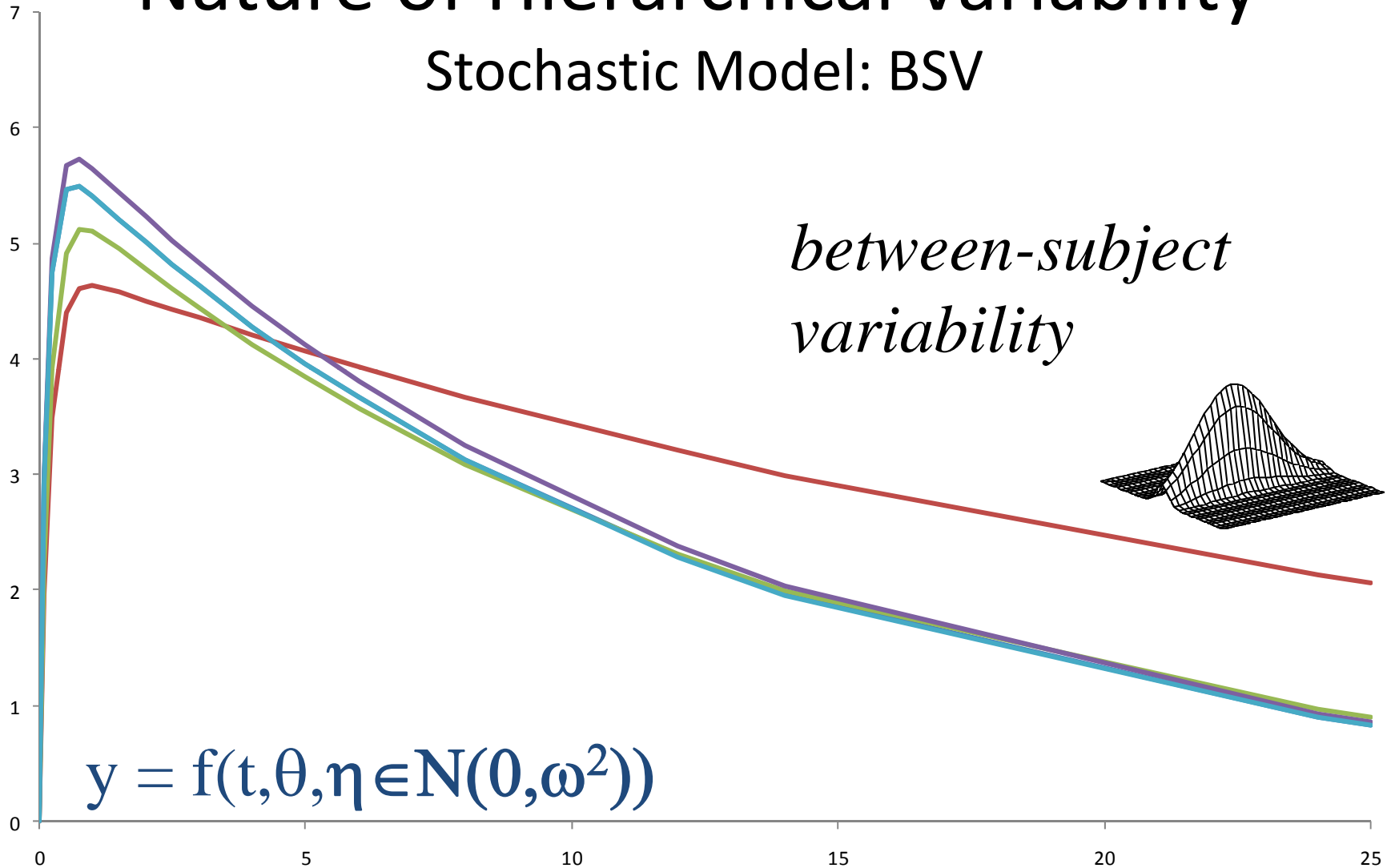
- Statistical-mathematical expression relating subject-specific PK or PD model parameter values to random variables describing population variability [*random effects*], e.g.:

$$CL = \theta + \eta$$

- Parameters:
 - CL, individualized value of drug clearance in the PK model
 - θ , population clearance
 - η , random effect, usually drawn from a normal (=Gaussian) density with a mean of zero and a variance of ω^2 :
 $\eta \sim \mathbf{N}(\mathbf{0}, \omega^2)$

Nature of Hierarchical Variability

Stochastic Model: BSV



A Note on Parameter Distributions

- Note: Density = Distribution (synonyms)
- Typically, variability in physiologic processes and spaces is reflected by:
 - Constrained (truncated) normal, or
 - Log-normal distributions,as negative values are unfeasible for such processes
- Thus, even if a parameter is e.g. log-normal, its natural log is more approximately normal:

$$CL = \theta \times \exp(\eta) \Rightarrow \log(CL) = \log(\theta) + \eta$$

PK/PD Parameter Probability Density

- Not necessarily restricted to Normal:

- Normal Distribution:

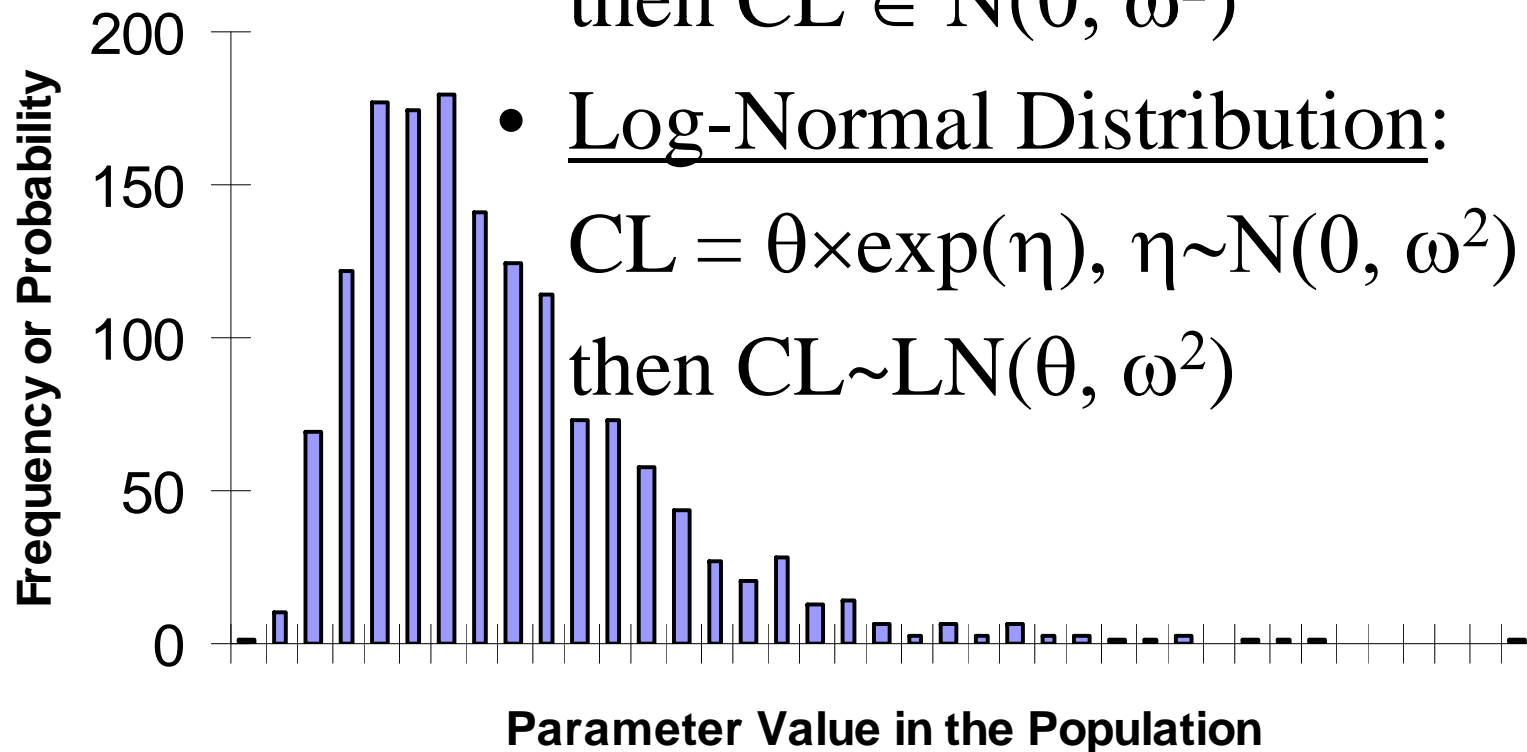
$$CL = \theta + \eta, \eta \sim N(0, \omega^2)$$

$$\text{then } CL \in N(\theta, \omega^2)$$

- Log-Normal Distribution:

$$CL = \theta \times \exp(\eta), \eta \sim N(0, \omega^2)$$

$$\text{then } CL \sim LN(\theta, \omega^2)$$

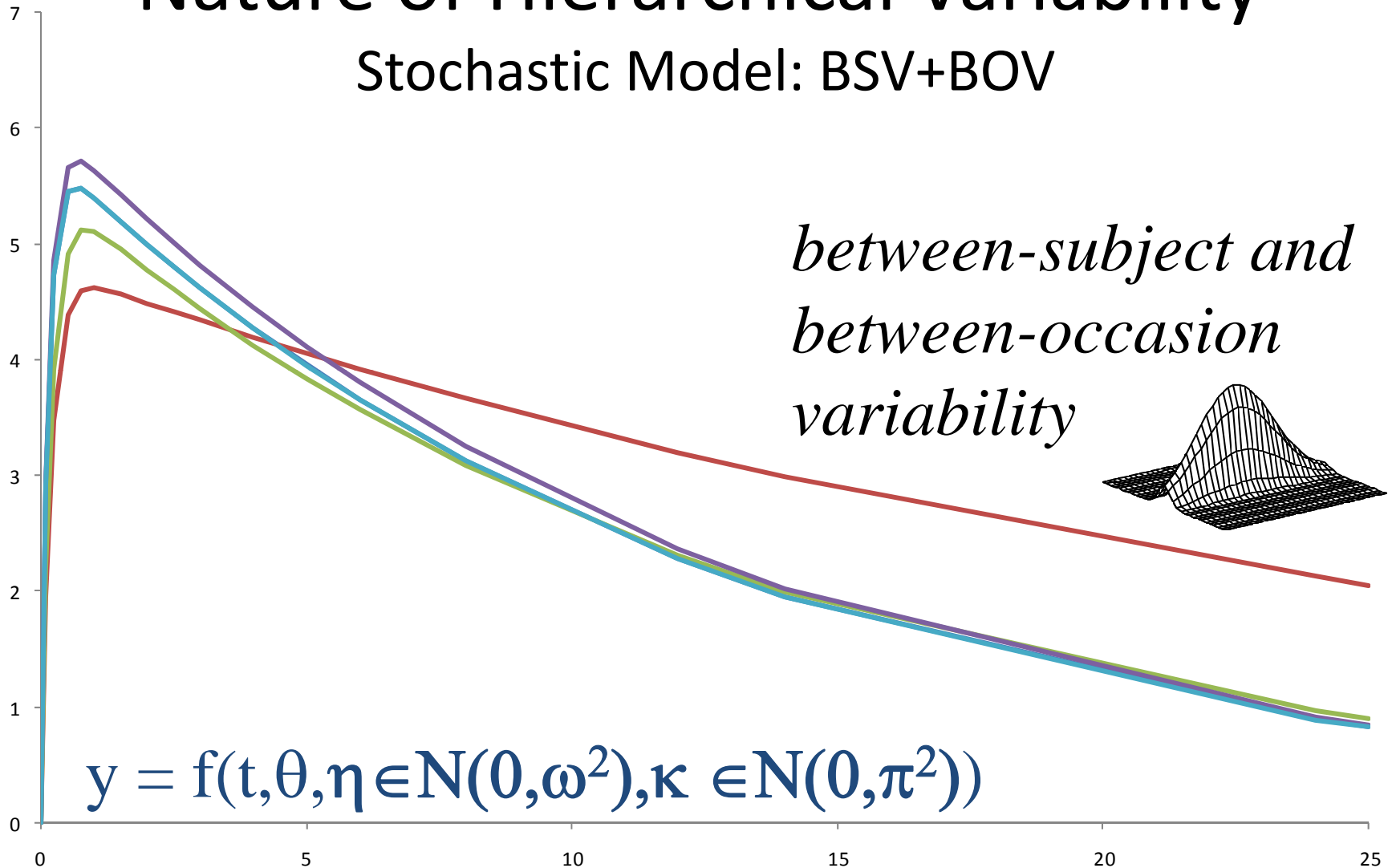


Between-Occasion Variability

- Statistical-mathematical expression relating subject-specific PK or PD model parameter values to random variables describing their variation between different study sessions
- Parameters: $CL = [\theta + \kappa] + \eta$
 - CL, value of drug clearance in the PK model
 - θ , slope relating clearance to body weight
 - η , BSV random effect, usually $\eta \in \mathbf{N}(\mathbf{0}, \omega^2)$
 - κ , BOV random effect, usually $\kappa \in \mathbf{N}(\mathbf{0}, \pi^2)$

Nature of Hierarchical Variability

Stochastic Model: BSV+BOV



Residual Unknown Variability

- Statistical-mathematical expression relating PK/PD or disease model predictions to the actual trial measurements (observations):

$$Y(\text{time}) = F(\text{time}) + \varepsilon(\text{time})$$

- In all generality:
 - Y , value of the observation (PK, PD, disease) at a certain time
 - F , corresponding value of the model prediction
 - ε , additive random variable describing the observation randomness not explained by the model
- Usually, $\varepsilon \in N(0, \sigma^2)$, where σ^2 may or may not be a function of time

Different Models for RUV

- Additive:

$$Y(\text{time}) = F(\text{time}) + \varepsilon(\text{time})$$

- Proportional:

$$Y(\text{time}) = F(\text{time}) \times [1 + \varepsilon(\text{time})]$$

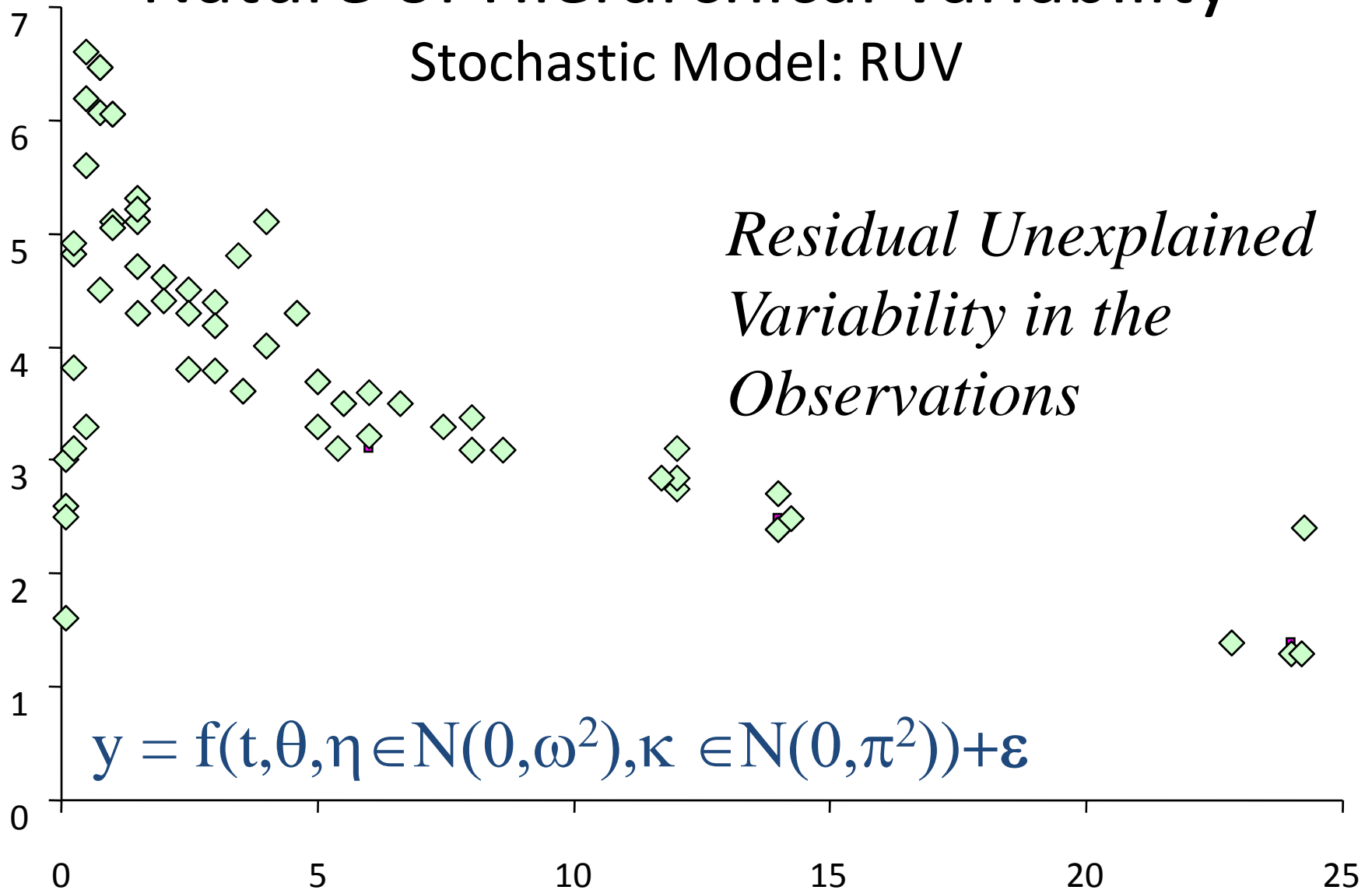
- Additive + proportional:

$$Y(\text{time}) = \varepsilon_1(\text{time}) + F(\text{time}) \times [1 + \varepsilon_2(\text{time})]$$

Nature of Hierarchical Variability

Stochastic Model: RUV

*Residual Unexplained
Variability in the
Observations*



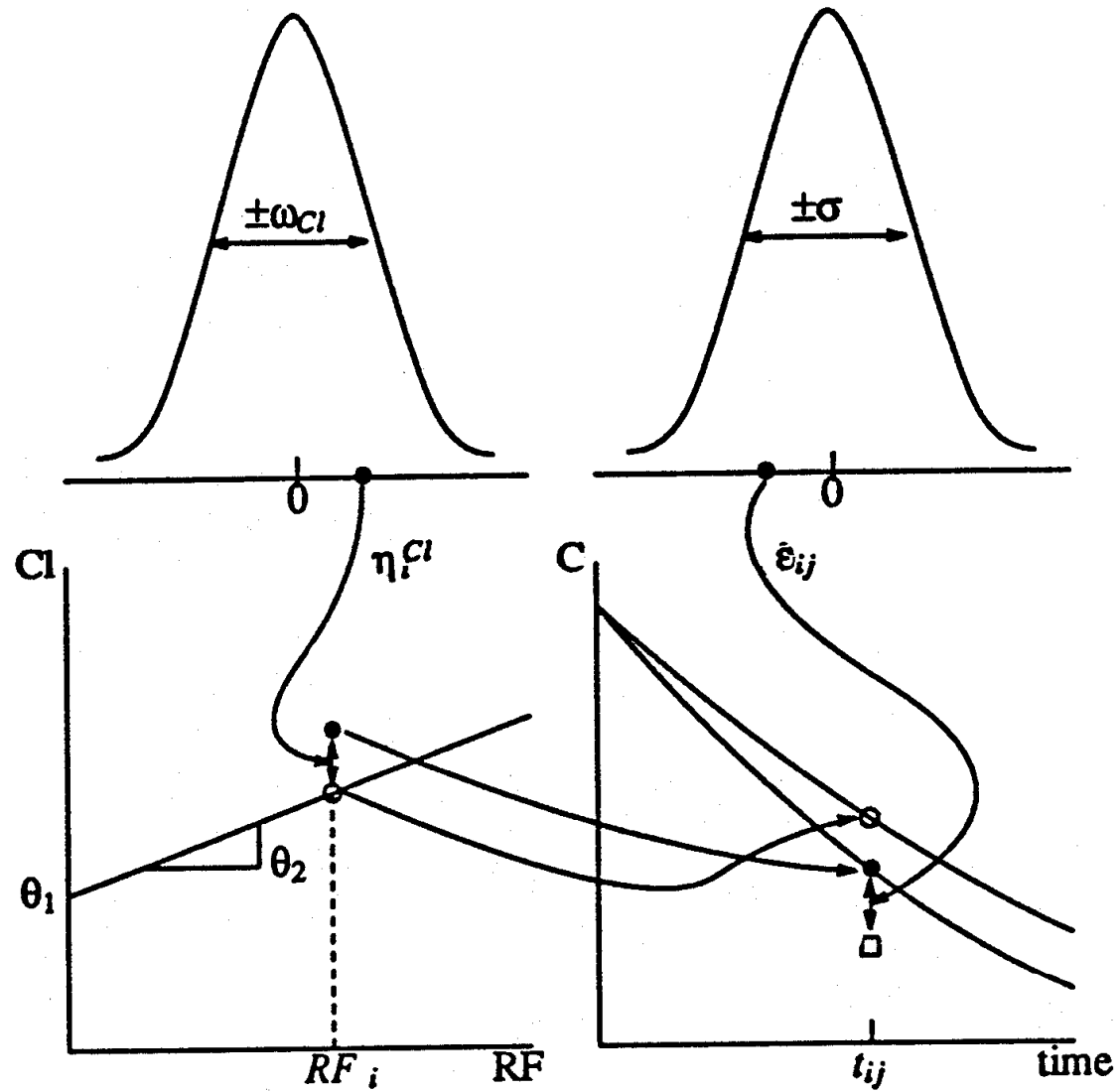
Covariate Model

- Mathematical expression relating individual-specific covariates (= explanatory variables) to typical values of the PK or PD model parameters, via other unknown parameters [*fixed effects*]: e.g. for paracetamol:

$$CL = \theta_1 + \theta_2 \times RF$$

- Parameters:
 - CL, typical value of drug clearance in the PK model
 - BW, individually-measured body weight
 - θ , slope (fixed effect) relating clearance to body weight

Population PK Variability



Bayes Theorem

Produce inferences about model parameters (θ) given data (y) and prior knowledge about model parameters ($p(\theta)$).

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \propto p(\theta)p(y|\theta)$$

Posterior = Prior \times Likelihood

Bayes Theorem (3rd form) and the “Harvard Question”

$$p(dz | +) = \frac{p(dz)}{p(dz) + \frac{p(+ | \sim dz)}{p(+ | dz)} \times p(\sim dz)}$$

Question:

A test for a disease has two outcomes, positive and negative (presence/absence of disease)

False negative rate (i.e., rate of giving a negative answer if the patient has a disease) is zero

False positive rate is 5% (i.e., rate of showing a positive result if the patient does not have the disease)

The Disease incidence is .1% (1 in 1000)

If the test is used in a randomly selected person and the person has a positive result

What is the probability that the person has the disease?

Individual and Population Info

# of subjects available # of data per subject	Many	Few
Many	Both individual and population information are robust	Individual information is most robust
Few	Population information is most robust	Neither individual and population information are robust