Multilevel Bayesian Analysis of Data
in the Presence of
Model Inadequacy and Measurement Error

Amir Shahmoradi

Center for Computational Oncology

Institute for Computational Engineering and Sciences

The University of Texas at Austin

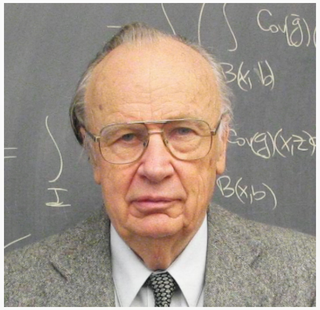
presented at

Babuska Forum

Nov 10 2017

Collaborators

J.T. Oden



Ivo Babuska

Alican Ozkan

Neda Ghousifam

Tom Yankeelov

Nichole Rylander

Financial Support

ICES
UT Austin
O'Donnell Foundation

US Department of Energy

US Department of Defense

National Science Foundation

American Physical Society

NASA

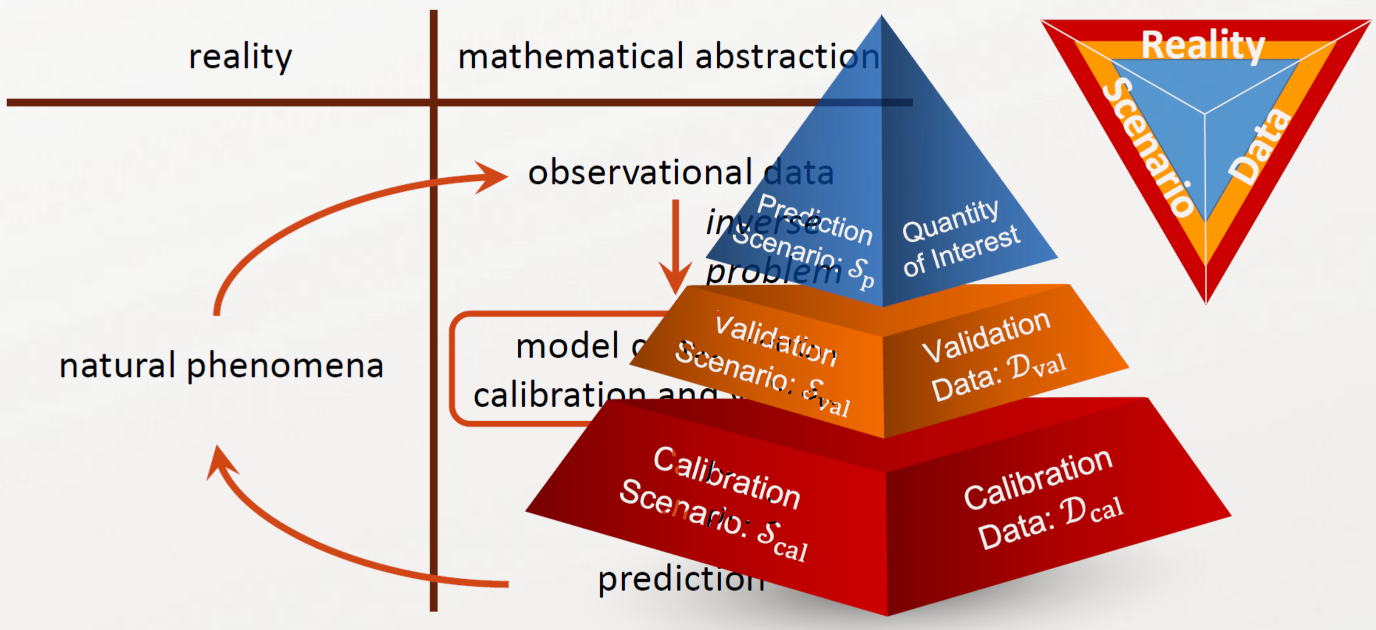
Probability Theory is the Logic of Science

How do we make a scientific prediction?

A very elementary depiction of the scientific method

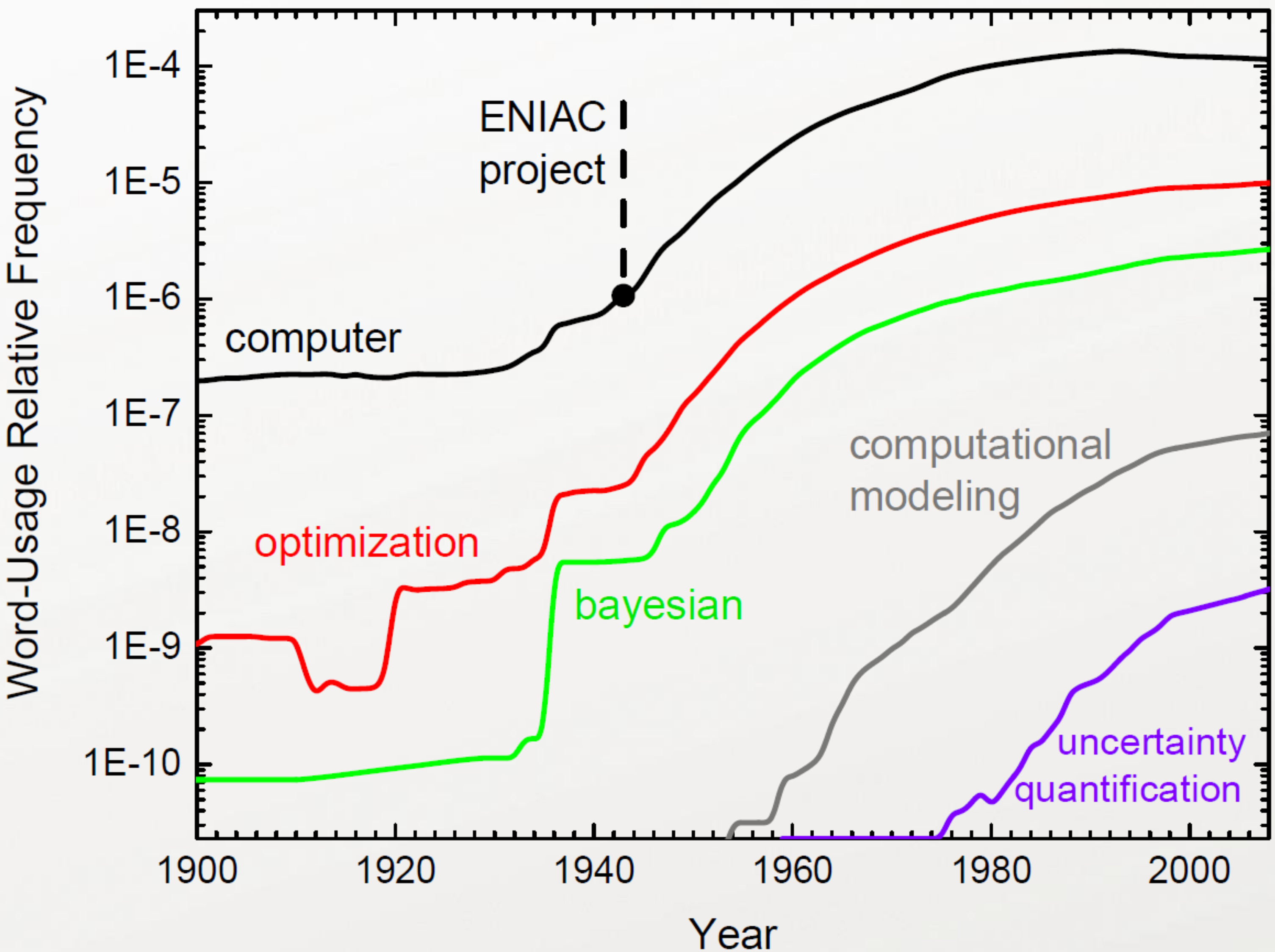
The prediction pyramid (tetrahedron)

(Oden et al., 2010, Computer predictions II)



Probability Theory is the Logic of Science

The field of uncertainty quantification is now 4 decades old.



Probability Theory is the Logic of Science

The field of uncertainty quantification is now 4 decades old.

Despite decades of UQ, wrong inference methods are ubiquitous in scientific literature.

Typical arguments against the correct scientific inference approaches:

Let the data speak for itself (frequentists)

I bet that would not make any difference (nominal Bayesians)

I need to survive (students, postdocs)

It's much more interesting to live not knowing than to have answers which might be wrong.

- *Richard Feynman*

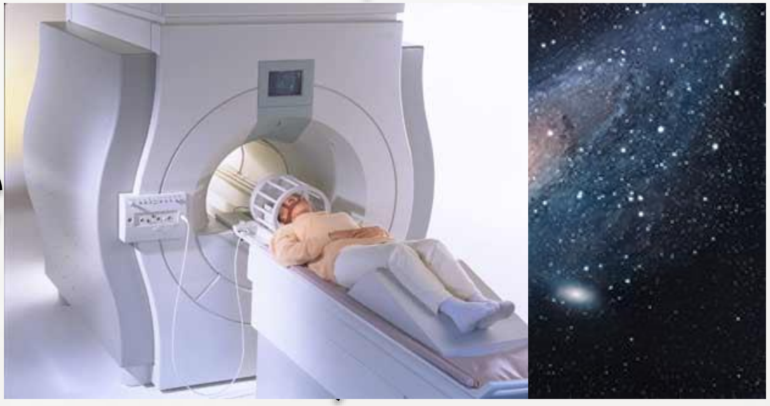
Probability Theory is the Logic of Science

The field of uncertainty quantification is now 4 decades old.

Despite decades of UQ, wrong inference methods are ubiquitous in scientific literature.

wrong answer

correct answer



academic survivability

mathematical complexity

practicality

practicality

It's much more interesting to live not knowing than to have answers which might be wrong.

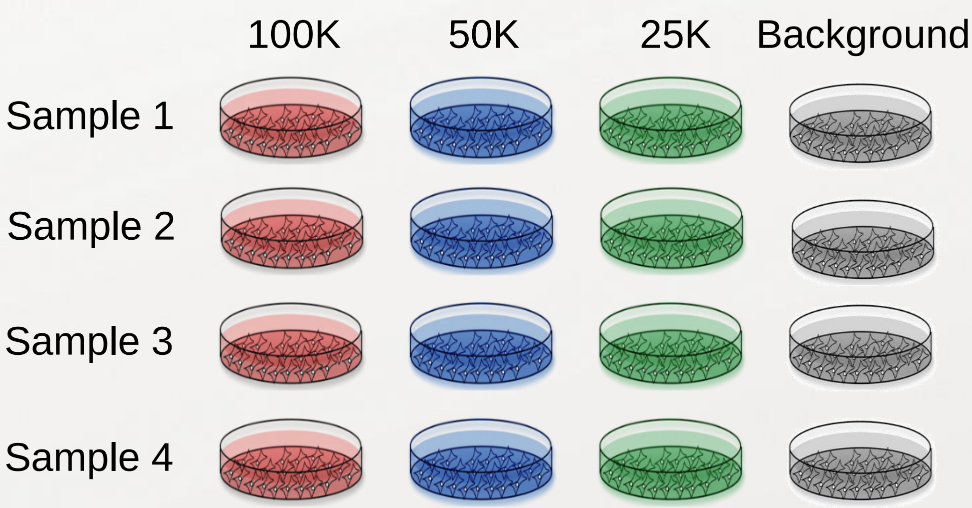
- *Richard Feynman*

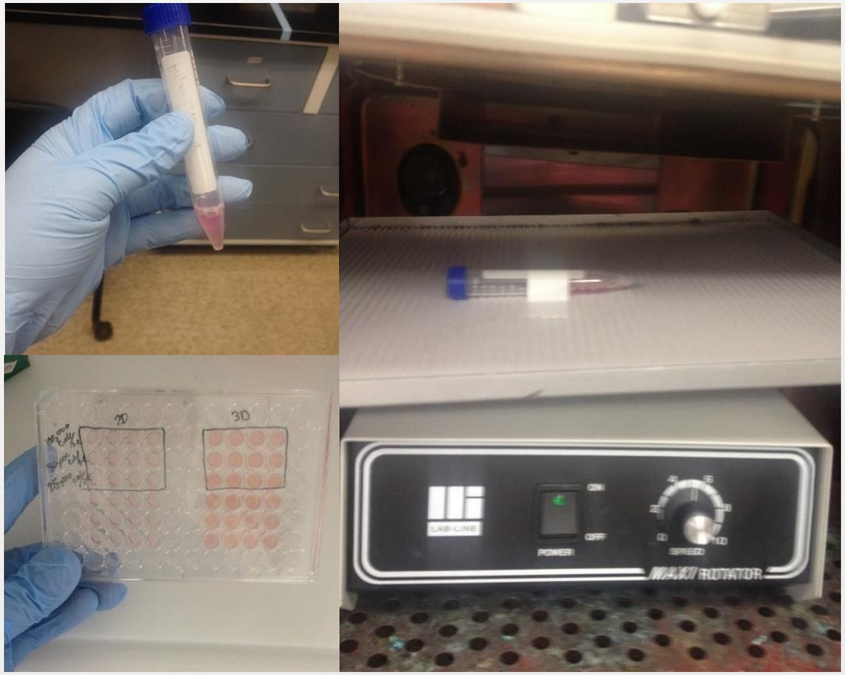
Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

Sample Preparation:

- Cells incubated with 5, 25, and 50 μg/ml Mitomycin-C at 37 ˚C for 30 minutes.
- Cell counting performed at: 0, 1, 3, 5, 7 days post sample preparation.
- Samples nutrients refreshed every 2 days.
- Three different cell densities: 100000, 50000, 25000 cells/ml, four samples per case,



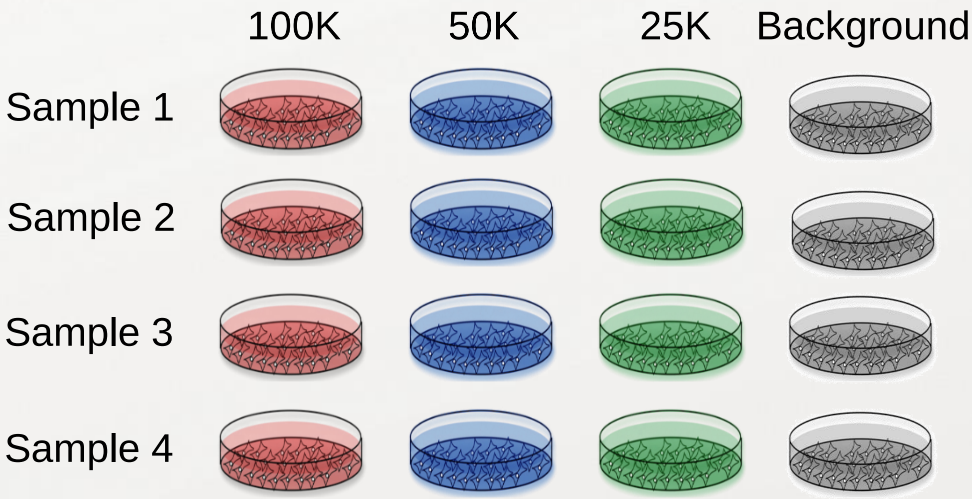


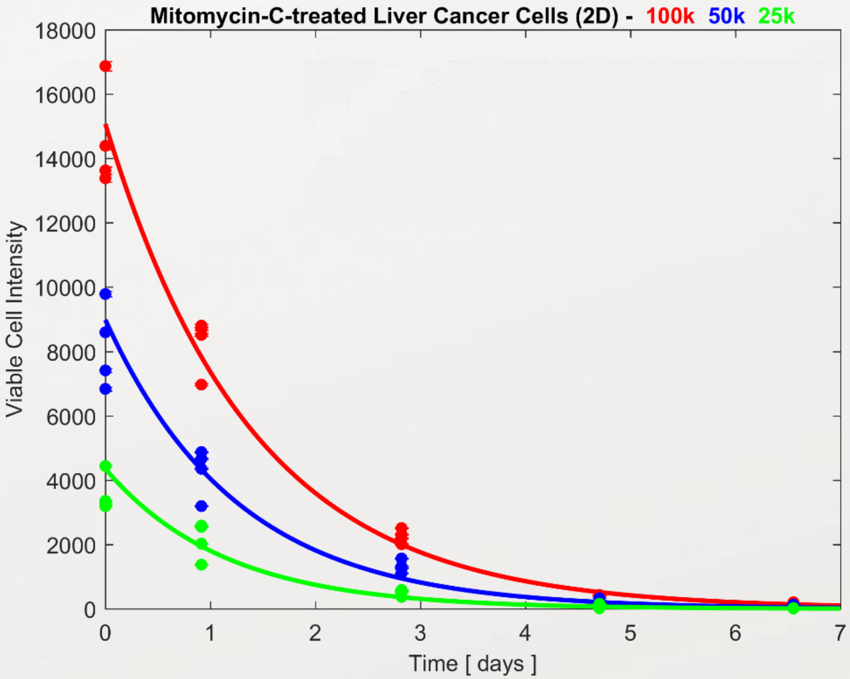
Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

Sample Preparation:

- Cells incubated with 5, 25, and 50 µg/ml Mitomycin-C at 37 ˚C for 30 minutes.
- Cell counting performed at: 0, 1, 3, 5, 7 days post sample preparation.
- Samples nutrients refreshed every 2 days.
- Three different cell densities: 100000, 50000, 25000 cells/ml, four samples per case,





Traditional orthodox solutions can lead to logical paradoxes

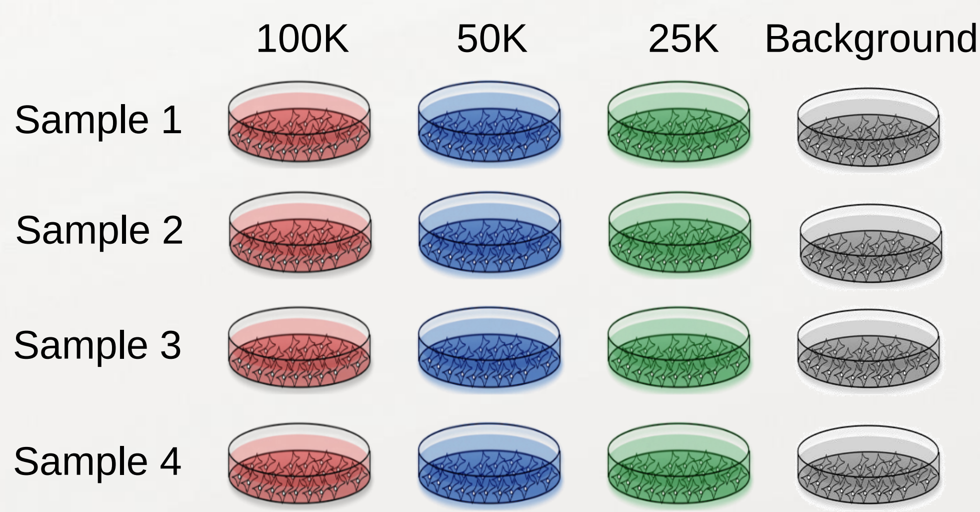
Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

$$I_{\text{obs}} \quad : \quad \text{observed fluorescence intensity [RFU]}$$
$$I_{\text{tru}} \quad : \quad \text{tumor cells intensity [RFU]}$$
$$I_{\text{bac}} \quad : \quad \text{background intensity [RFU]}$$

$$I_{\text{obs}} = I_{\text{tru}} + I_{\text{bac}}$$
$$\Rightarrow \quad \widehat{I}_{\text{tru}} = I_{\text{obs}} - \widehat{I}_{\text{bac}}$$

$\widehat{I}_{\text{tru}}$ [RFU]

day 0			
concentration	**100K**	**50K**	**25K**
sample 1	13628.8	7411	3292
sample 2	16872.8	8595	3353
sample 3	13383.8	6836	3190
sample 4	14390.8	9786	4439



I_{obs} [RFU]

day 0				
concentration	**100K**	**50K**	**25K**	**BG**
sample 1	32928	26710	22591	19054
sample 2	36172	27894	22652	19390
sample 3	32683	26135	22489	19343
sample 4	33690	29085	23738	19410

Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

$I_{\rm obs}$: observed fluorescence intensity [RFU]

$I_{\rm tru}$: tumor cells intensity [RFU]

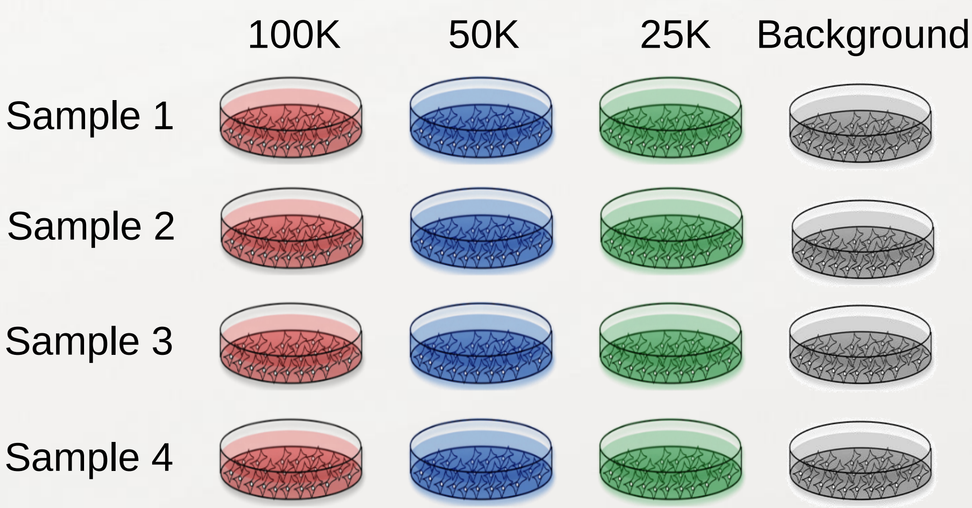
$I_{\rm bac}$: background intensity [RFU]

$$I_{\rm obs} = I_{\rm tru} + I_{\rm bac}$$

$$\Rightarrow \quad \widehat{I}_{\rm tru} = I_{\rm obs} - \widehat{I}_{\rm bac}$$

$\widehat{I}_{\rm tru}$ [RFU]

day 1			
concentration	**100K**	**50K**	**25K**
sample 1	8706	4353	1373
sample 2	6971	4873	2025
sample 3	8516	3193	2588
sample 4	8803	4664	2547



$I_{\rm obs}$ [RFU]

day 1				
concentration	**100K**	**50K**	**25K**	**BG**
sample 1	28213	23860	20880	19407
sample 2	26478	24380	21532	19464
sample 3	28023	22700	22095	19605
sample 4	28310	24171	22054	19552

Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

I_{obs} : observed fluorescence intensity [RFU]

I_{tru} : tumor cells intensity [RFU]

I_{bac} : background intensity [RFU]

$$I_{\mathrm{obs}} = I_{\mathrm{tru}} + I_{\mathrm{bac}}$$

$$\Rightarrow \quad \widehat{I}_{\mathrm{tru}} = I_{\mathrm{obs}} - \widehat{I}_{\mathrm{bac}}$$

$\widehat{I}_{\mathrm{tru}}$ [RFU]

day 3			
concentration	**100K**	**50K**	**25K**
sample 1	2008.75	1317	371.8
sample 2	2195.75	1565	585.8
sample 3	2309.75	1117	509.8
sample 4	2509.75	1253	560.8

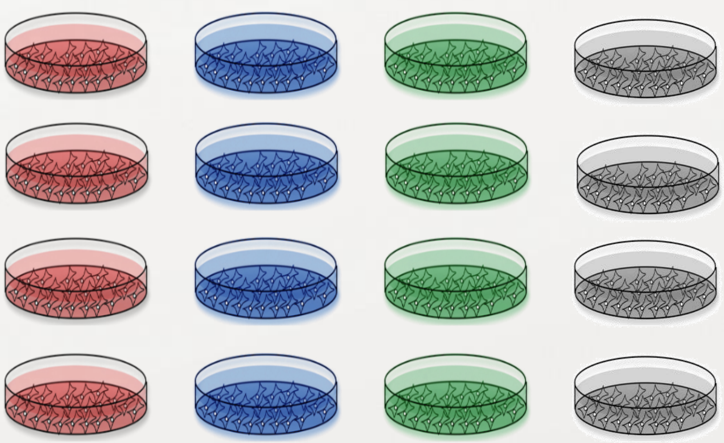
100K 50K 25K Background

Sample 1

Sample 2

Sample 3

Sample 4



I_{obs} [RFU]

day 3				
concentration	**100K**	**50K**	**25K**	**BG**
sample 1	6469	5777	4832	4442
sample 2	6656	6025	5046	4442
sample 3	6770	5577	4970	4460
sample 4	6970	5713	5021	4497

Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

I_{obs} : observed fluorescence intensity [RFU]

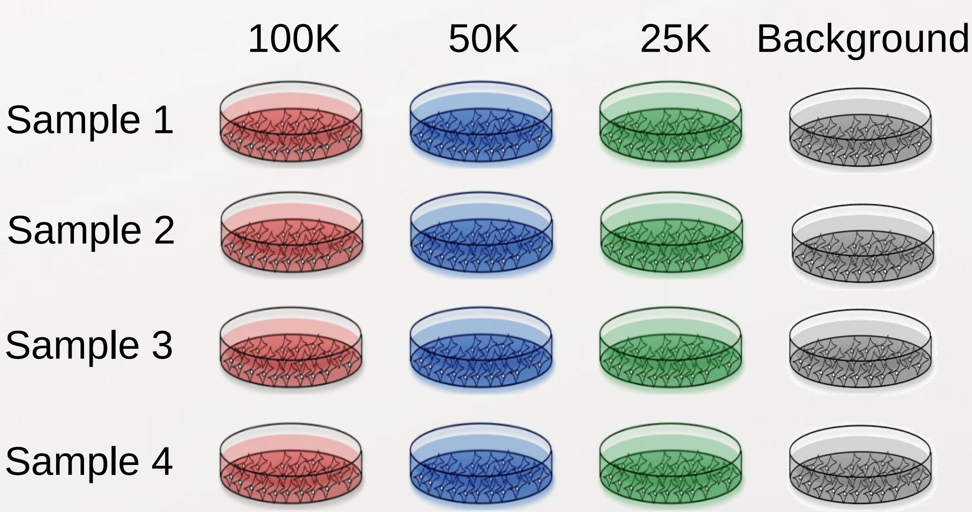
I_{tru} : tumor cells intensity [RFU]

I_{bac} : background intensity [RFU]

$$
\begin{aligned}
& I_{\text{obs}} = I_{\text{tru}} + I_{\text{bac}} \\
\Rightarrow \quad & \widehat{I}_{\text{tru}} = I_{\text{obs}} - \widehat{I}_{\text{bac}}
\end{aligned}
$$

$\widehat{I}_{\text{tru}}$ [RFU]

day 5			
concentration	**100K**	**50K**	**25K**
sample 1	380.25	62.25	15.25
sample 2	409.25	178.3	54.25
sample 3	432.25	248.3	118.3
sample 4	397.25	327.3	152.3



I_{obs} [RFU]

day 5				
concentration	**100K**	**50K**	**25K**	**BG**
sample 1	3500	3182	3135	3132
sample 2	3529	3298	3174	3126
sample 3	3552	3368	3238	3087
sample 4	3517	3447	3272	3134

Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

I_{obs} : observed fluorescence intensity [RFU]

I_{tru} : tumor cells intensity [RFU]

I_{bac} : background intensity [RFU]

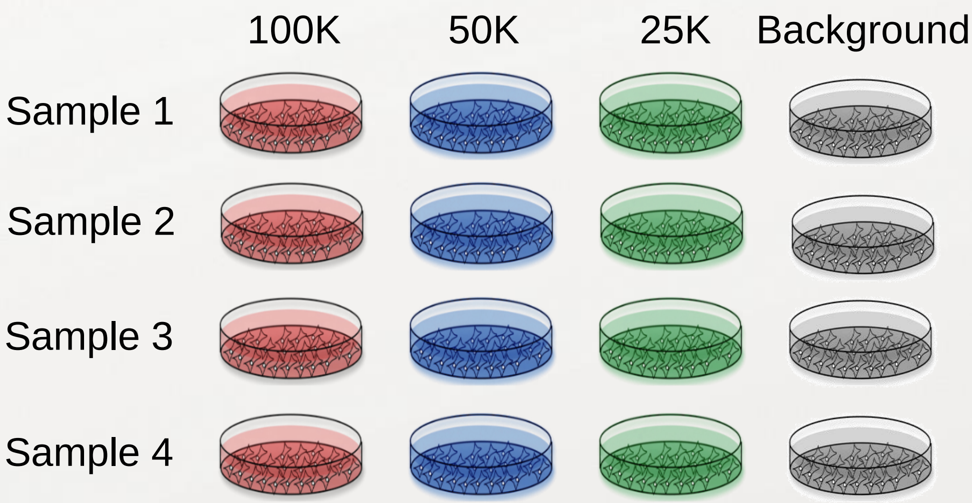
$$I_{\mathrm{obs}} = I_{\mathrm{tru}} + I_{\mathrm{bac}}$$

$$\Rightarrow \quad \widehat{I}_{\mathrm{tru}} = I_{\mathrm{obs}} - \widehat{I}_{\mathrm{bac}}$$

$$\leq \mathbf{0}$$

$\widehat{I}_{\mathrm{tru}}$ [RFU]

day 7			
concentration	**100K**	**50K**	**25K**
sample 1	108.75	95.75	**-64.3**
sample 2	202.75	**-58.3**	**-35.3**
sample 3	160.75	44.75	**-12.3**
sample 4	166.75	126.8	24.75



I_{obs} [RFU]

day 7				
concentration	**100K**	**50K**	**25K**	**BG**
sample 1	2592	2579	2419	2476
sample 2	2686	2425	2448	2483
sample 3	2644	2528	2471	2493
sample 4	2650	2610	2508	2481

Traditional orthodox solutions can lead to logical paradoxes

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

I_{obs} : observed fluorescence intensity [RFU]

I_{tru} : tumor cells intensity [RFU]

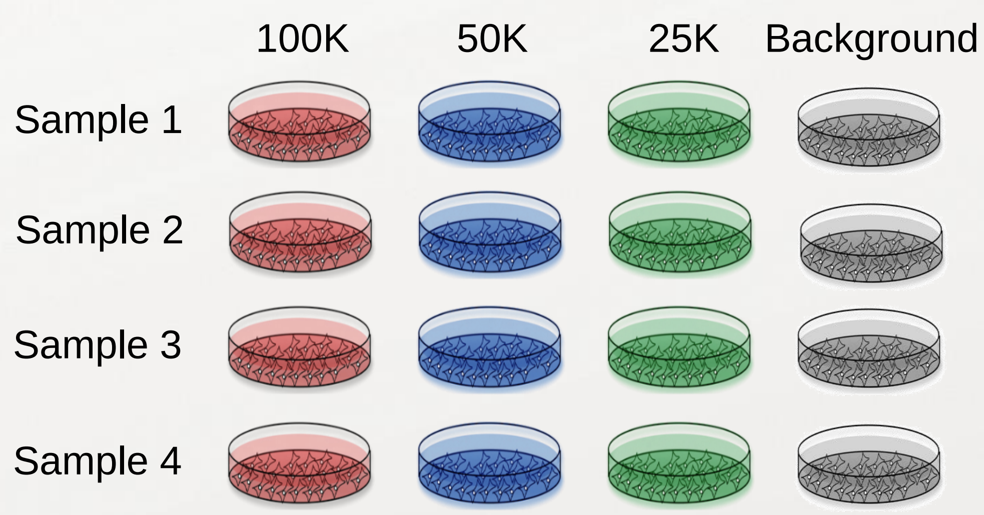
I_{bac} : background intensity [RFU]

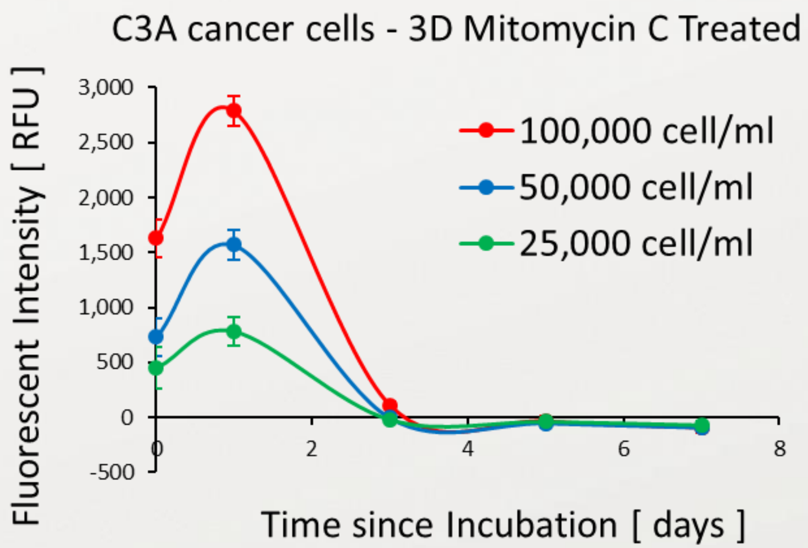
$$I_{\text{obs}} = I_{\text{tru}} + I_{\text{bac}}$$

$$\Rightarrow \quad \widehat{I}_{\text{tru}} = I_{\text{obs}} - \widehat{I}_{\text{bac}} \leq \mathbf{0}$$

$\widehat{I}_{\text{tru}}$ [RFU]

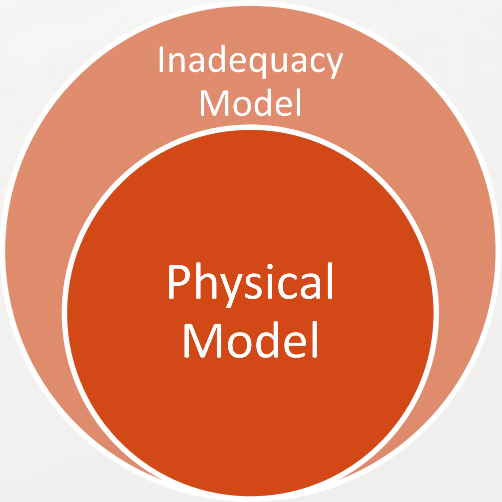
day 7				
concentration	**100K**	**50K**	**25K**	**1σ error**
sample 1	**-46.3**	**-126**	**-59.3**	±7.1
sample 2	**-59.3**	**-41.3**	**-58.3**	±7.1
sample 3	**-157**	**-119**	**-114**	±7.1
sample 4	**-96.3**	**-70.3**	**-45.3**	±7.1

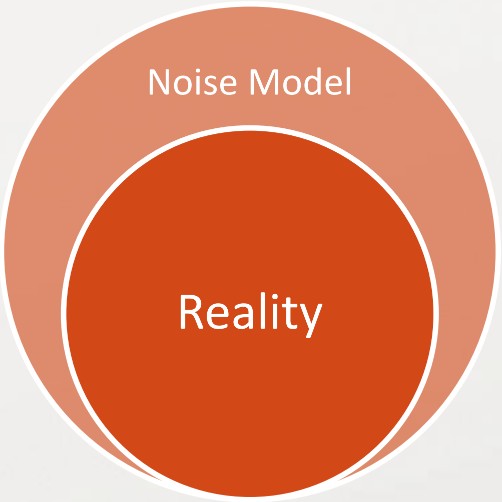




Outline

- Bayesian vs. Frequentist inference
- Two types of uncertainty with regards to experimental data:
  - Model inadequacy
  - Experimental noise
- Wrong choice of inadequacy can lead to
  - **Logical paradoxes**: negative number of tumor cells, negative concentration, …
  - Increased likelihood of **false-negative** and **false-positive** conclusions
- Confusion of noise model with model inadequacy can lead to:
  - **Logical paradoxes**: zero likelihood
  - Increased likelihood of **false-negative** and **false-positive** conclusions
- Stochastic integration techniques for model selection (Parallel Tempering, Lebesgue Monte Carlo)





Digression: The Four of Classes of Probability Definition

(Shahmoradi et al 2017)

Interpretation	Classical	Frequentist	Bayesian	Propensity
Definition of probability	relative frequency	long-run relative frequency of occurrence	degree of belief	tendency of occurrence
Philosophical basis	principle of indifference	infinite repetition of identical experiments	prior belief + current evidence	probability from single experiment
Domain of definition	physical events, hypotheses	physical events	physical events, hypotheses	physical events
Epistemology†	objective	objective	objective, subjective	objective
Primary field of origin	Astronomy, Physics, Economics	Philosophy, Biology, Statistics	Astronomy, Physics, Economics	Philosophy
Major advocates & contributors	J. Bernoulli (1713) [9], de Moivre (1718) [39], Laplace (1774) [108, 163], D. Bernoulli (1777) [8], Condorcet (1785) [36], Lubbock (1830) [118], Poisson (1837) [35, 142], de Morgan (1838) [130], Maxwell (1850) [21] Jevons (1874) [99], Boltzmann (1877) [13], Gibbs (1902) [69], Neyman (1937) [133]	Cournot (1843) [27], Ellis (1843) [52], Boole (1854) [14], Venn (1866) [168], Chrystal (1889) [77], Galton (1891) [19], Fisher (1921) [57], von Mises (1931) [169], Reichenbach (1934) [151], Popper (1934) [146], Pearson†† (1941) [140], Kendall (1943) [100], Cramér (1946) [31], Neyman (1950) [134], Feller (1950) [53]	Bayes (1763) [5], Laplace (1774) [108, 163], de Morgan (1838) [40], Poincaré (1912) [141], Keynes (1921) [102], Ramsey (1931) [150], de Finetti (1937) [37], Jeffreys (1939) [98], R.T. Cox (1946) [28, 29], Carnap (1950) [23], Good (1950) [76], Pólya (1954) [145], Savage (1954) [158], Zellner (1971) [179], Jaynes (2003) [95], Jeffrey (2004) [97], Lindley (2006) [115]	Popper (1959) [147], Gillies (1973) [72, 73], Fetzer (1974) [54, 55], Miller (1994) [128]
Major limitations	reliance on the principle of indifference;	reliance on repeated identical experiments (not applicable to single non-recurring events); indeterminate mathematical definition; reference class problem;	computationally challenging;	vague definition; incoherence (e.g., Humphreys' Paradox, see Section 2.1.3); reference class problem;
Remarks	special case of Bayesian probability; also known as *a priori probability*;	also known as *chance*, *direct probability*, *physical probability*, *objective probability*, *empirical probability*, *statistical probability*, *a posteriori probability*;	also known as *credibility*, *logical probability*, *inverse probability*, *inductive probability*, *epistemic probability*, *subjective probability*, *evidential probability*;†††	proposed to circumvent the limitations of frequentist interpretation;

Digression: Two Philosophically distinct approaches to statistical inference

Frequentist Inference

Neyman–Pearson–Wald theory

Jerzy Neyman (1894 – 1981)
Statistician, Astronomer

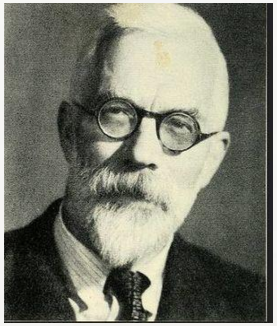
Known for
- Hypothesis testing
- Statistics of galaxy clusters

Egon Pearson (1895 – 1980)
Statistician

Known for
- Hypothesis testing
- Karl Pearson’s son

Abraham Wald (1902 – 1950)
Statistician

Known for
- Hypothesis testing
- Neyman–Pearson–Wald theory

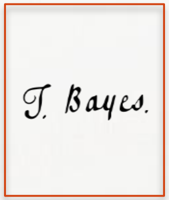


Ronald Fisher
(1890 – 1962)
Bio-statistician

Fiducial Inference
Maximum Likelihood

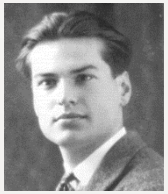
Bayesian Inference

Bayesian probability theory

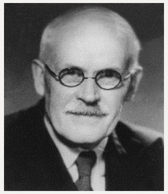


Thomas Bayes
(1894 – 1981)
Statistician

Pierre Laplace
(1749 – 1827)
Astronomer / Mathematician



Bruno de Finetti
(1906 – 1985)
Statistician



Harold Jeffreys
(1891 – 1989)
Astronomer / Geophysicist

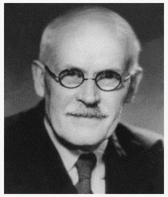
Richard Cox
(1898 – 1991)
Physicist

Edwin Jaynes
(1922 – 1998)
Physicist

Digression: Two Philosophically distinct approaches to Bayesian inference

Objective Bayesian

Prior knowledge has an exact mathematical definition



Harold Jeffreys
(1891 – 1989)
Astronomer / Geophysicist

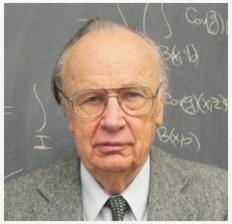
Edwin Jaynes
(1922 – 1998)
Physicist

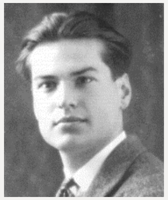
Richard Cox
(1898 – 1991)
Physicist

George Polya
(1887 – 1985)
Mathematician

Subjective Bayesian

Prior knowledge is anything I wish as long as I won't lose in betting





Bruno de Finetti
(1906 – 1985)
Statistician

Frank P. Ramsey
(1903 – 1930)
Philosopher

Rudolf Carnap
(1891 – 1970)
Philosopher

Leonard J. Savage
(1917 – 1971)
Statistician

Digression: There is only one type of uncertainty in the world – epistemic

Suppose you have **blurred** vision.
You throw a die **once**, and read your observation (possibly wrong reading).
What is the **type of uncertainty** in your observation?

Frequentist Inference

The uncertainty is due to my **lack of knowledge.**
I **can reduce uncertainty** with better vision**.**
Therefore, the **uncertainty** is **epistemic**.

Bayesian Inference

The uncertainty is due to my **lack of knowledge.**
I **can reduce uncertainty** with better vision**.**
Therefore, the **uncertainty** is **epistemic**.

Digression: There is only one type of uncertainty in the world – epistemic

Suppose you have **perfect** vision.
You throw a die **multiple times**, and read your observations.
What is the **type of uncertainty** in your observations?

Frequentist Inference

The uncertainty is **inherent in the experiment.**
I **cannot reduce** uncertainty any further**.**
Therefore, the **uncertainty** is **aleatoric**.

Bayesian Inference

The uncertainty is due to my **lack of knowledge:**

1. Wrong / **inadequate** model.
2. Lack of sufficiently-detailed data which leads to inadequate model.

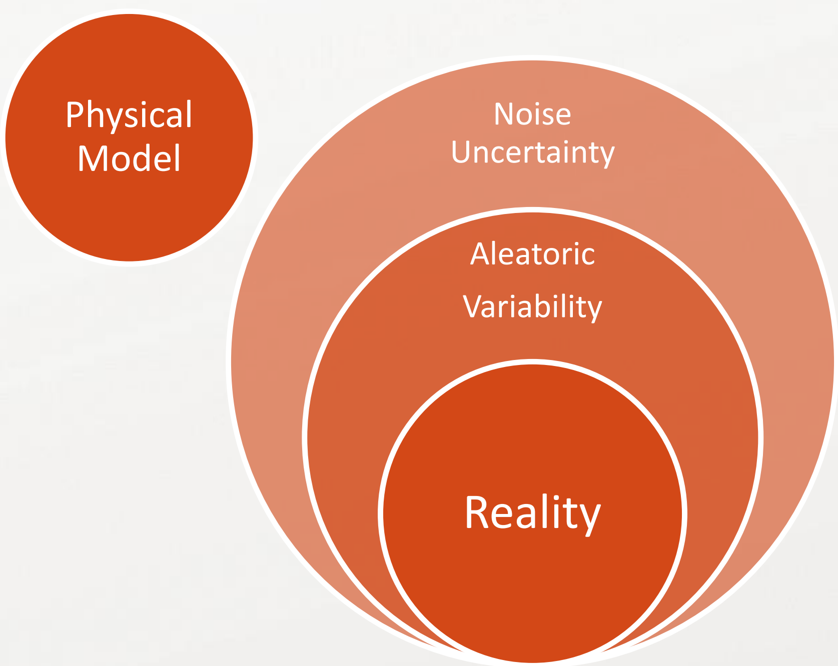
I **can reduce uncertainty** with better data / model**.**
Therefore, the **uncertainty** is **epistemic**.

Digression: There is only one type of uncertainty in the world – epistemic

Suppose you have **perfect** vision.
You throw a die **multiple times**, and read your observations.
What is the **type of uncertainty** in your observations?

Frequentist Inference

The uncertainty is **inherent in the experiment.**
I **cannot reduce** uncertainty any further**.**
Therefore, the **uncertainty** is **aleatoric**.

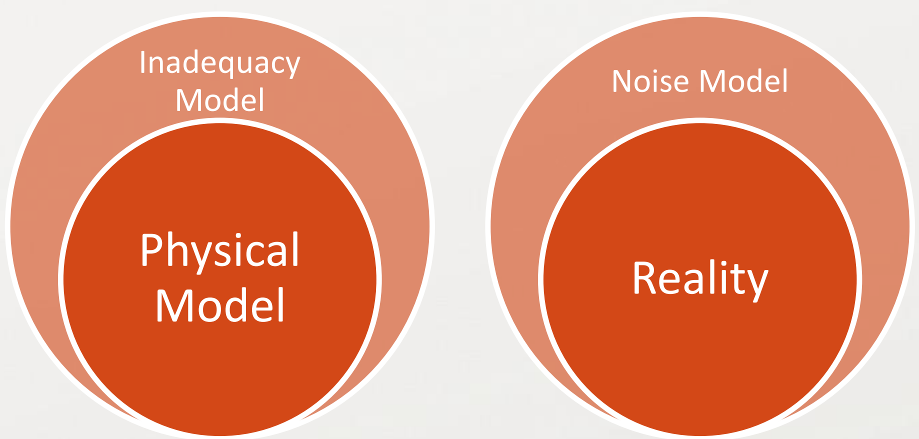


Bayesian Inference

The uncertainty is due to my **lack of knowledge:**

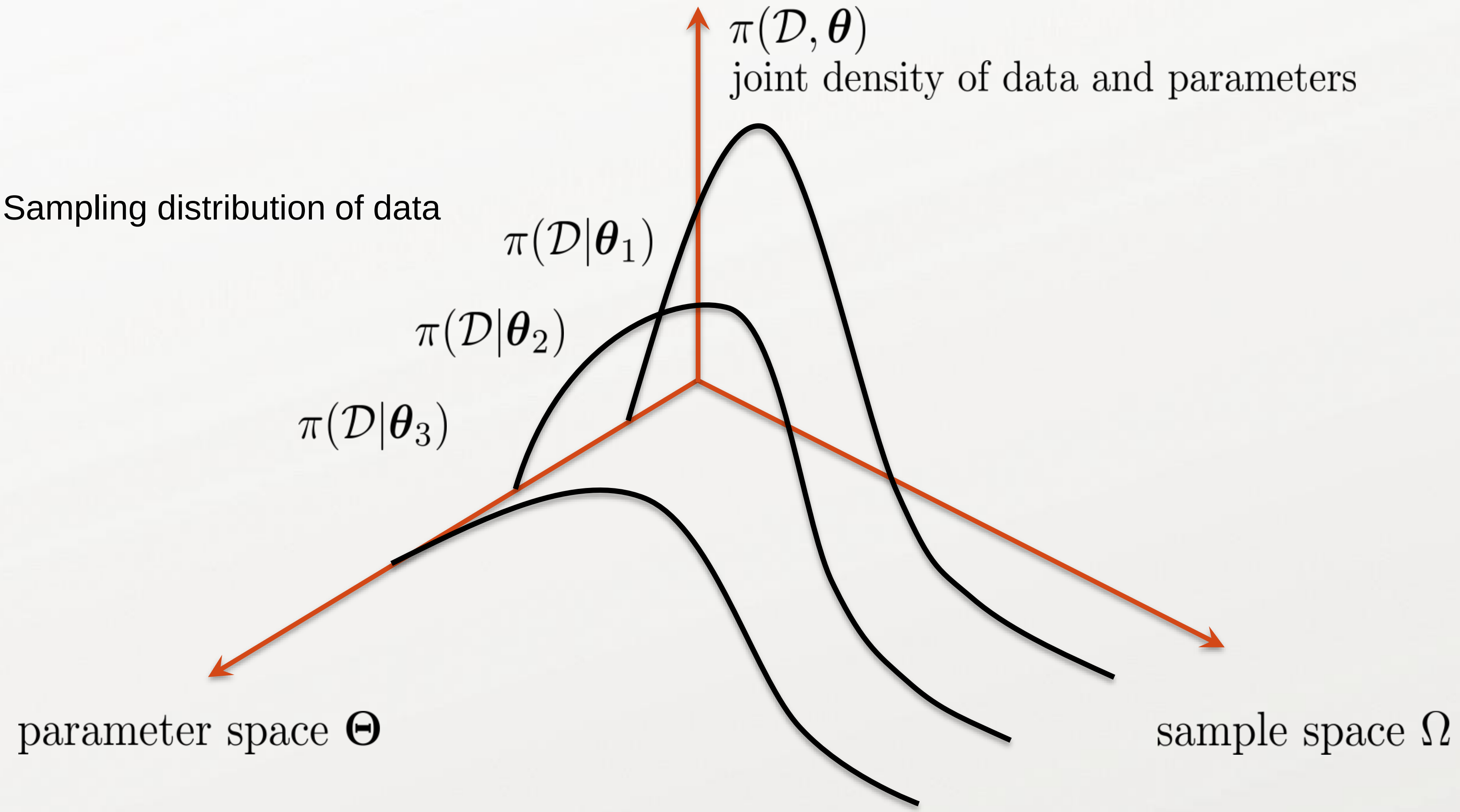
1. Wrong / **inadequate** model.
2. Lack of sufficiently-detailed data which leads to inadequate model.

I **can reduce uncertainty** with better data / model**.**
Therefore, the **uncertainty** is **epistemic**.



Digression: The world from a Frequentist perspective

Data is **random variable.**
Parameters are **fixed.**

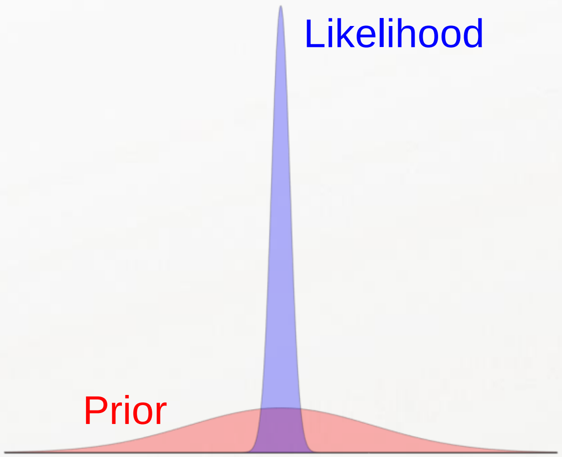


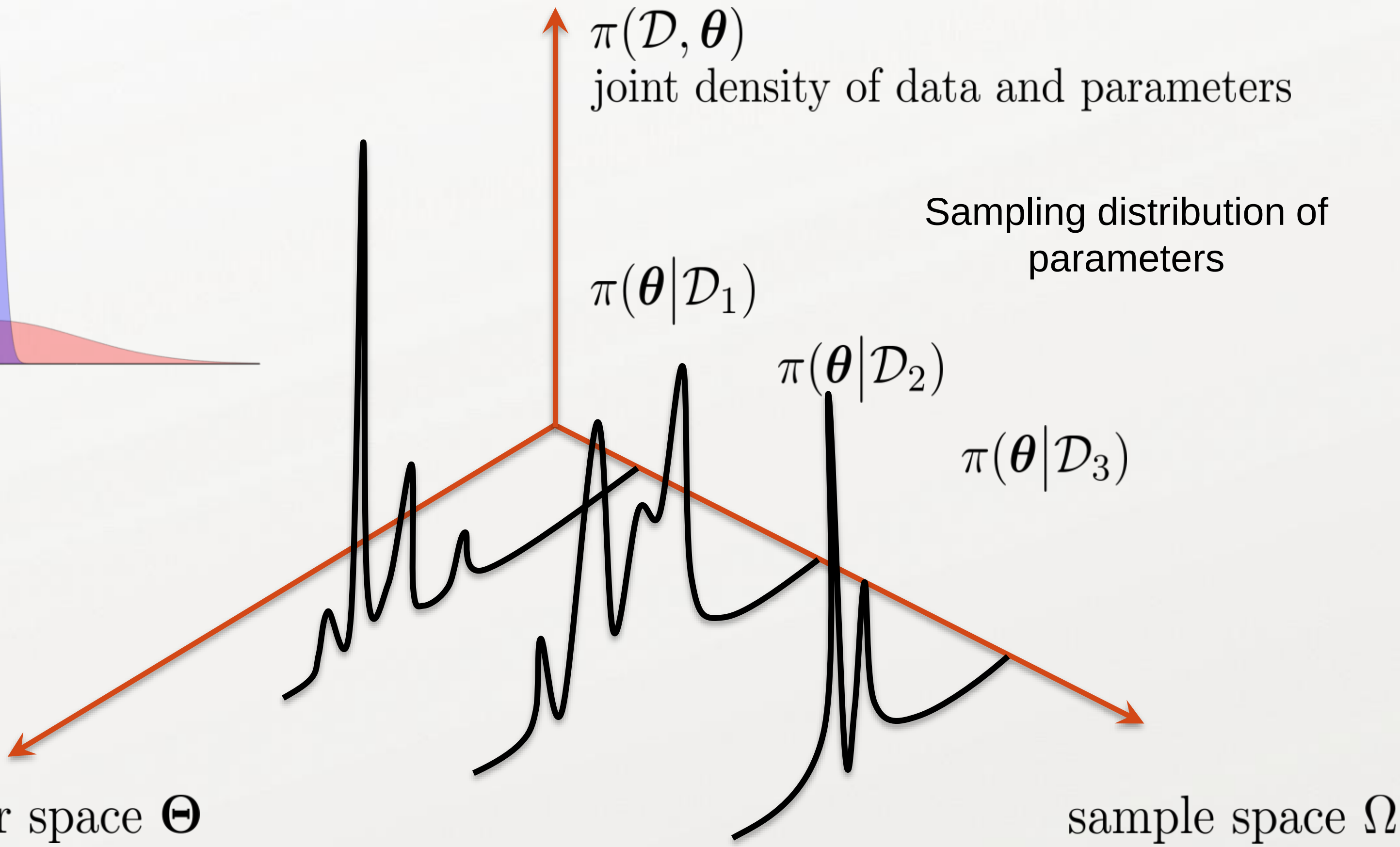
Digression: The world from a Bayesian perspective

Data is **fixed.**
Parameters are **random variables.**

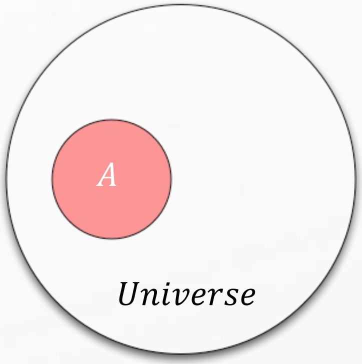
Bayes rule:

$$\pi(\boldsymbol{\theta}|\mathcal{D}, M) = \frac{\pi(\mathcal{D}|\boldsymbol{\theta}, M)\ \pi(\boldsymbol{\theta}|M)}{\pi(\mathcal{D}|M)}$$



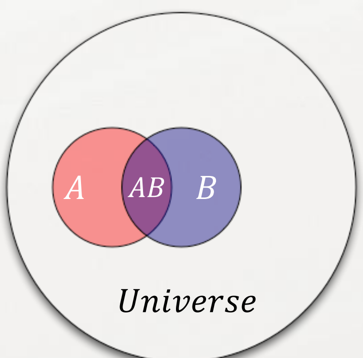


Digression: Bayes rule - The optimal method of updating knowledge



$$P(A) = \frac{A}{U}$$

$$P(B) = \frac{B}{U}$$



$$P(AB) = \frac{AB}{U}$$

$$P(A|B) = \frac{AB}{B} = \frac{\frac{AB}{U}}{\frac{B}{U}} = \frac{P(AB)}{P(B)}$$

$$P(B|A) = \frac{AB}{A} = \frac{\frac{AB}{U}}{\frac{A}{U}} = \frac{P(AB)}{P(A)}$$

Bayes rule

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Digression: Bayes rule - The optimal method of updating knowledge

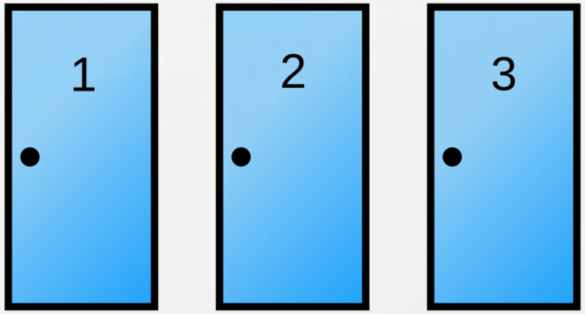
The Monty Hall Problem and Bayes Rule

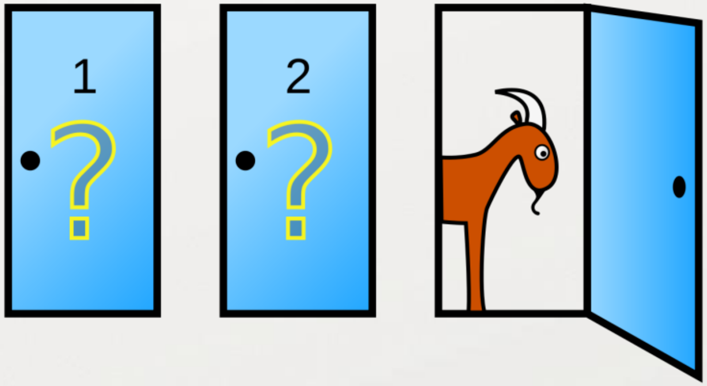
Steve Selvin, 1975, "A problem in probability (letter to the editor)". *American Statistician*

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the two others, goats. **You pick a door, say No. 1, and the host opens another door, say No. 3, which has a goat**. He then says to you, "Do you want to pick door No. 2?"

Question:
Is it to your advantage to switch your choice from door 1 to door 2?

Correct answer: The advantage of switching door, depends on your knowledge about the host’s decision.



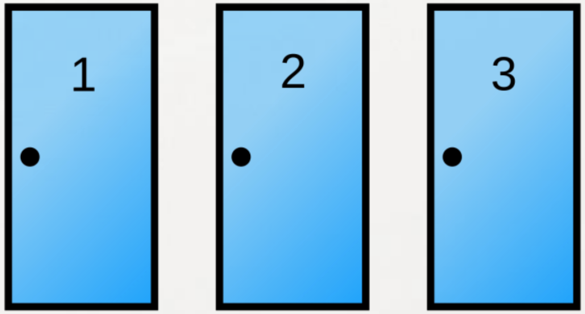


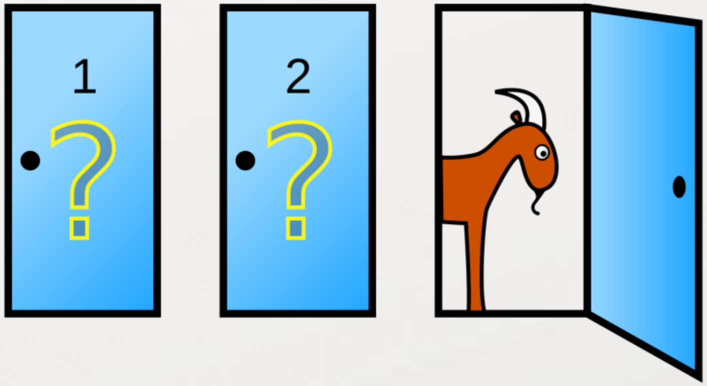
Digression: Bayes rule - The optimal method of updating knowledge

Case 1: Informed Host: Knows where the car is, and will not open the door that leads to car.

You, the guest, choose door 1. The **Informed** host **consciously** opens door 3.

$$P(C2|H3,G1) = \frac{P(H3|C2,G1)\ P(C2)}{P(H3|G1)}$$





Digression: Bayes rule - The optimal method of updating knowledge

Case 1: Informed Host: Knows where the car is, and will not open the door that leads to car.

You, the guest, choose door 1. The **Informed** host **consciously** opens door 3.

Prior knowledge about the car being behind door 2 (C2):

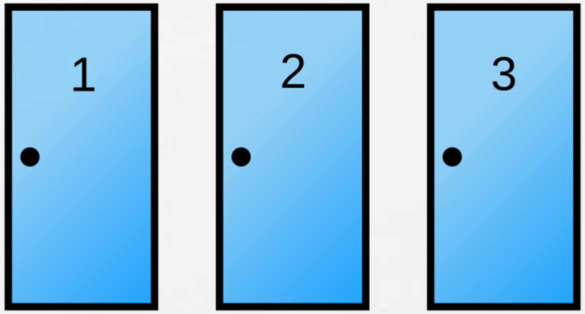
$$P(C2|G1) = P(C1|G1) = P(C3|G1) = P(C2) = \mathbf{\frac{1}{3}}$$

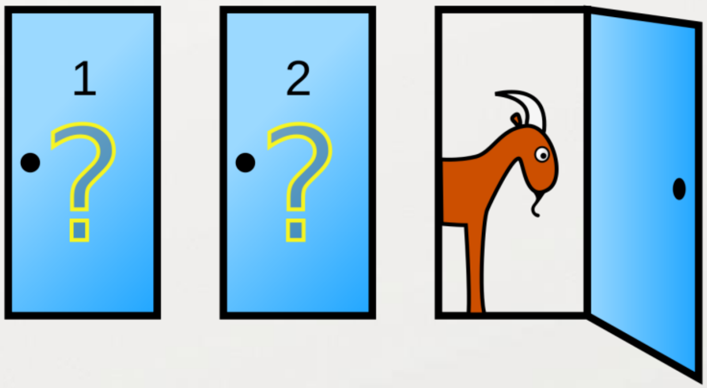
New knowledge: The informed host chooses door 3 (**H3**) (He knows that the car is behind door 2)

$$P(H3|C2, G1) = \mathbf{1} \qquad\qquad P(H3|G1) = \mathbf{\frac{1}{2}}$$

Update your knowledge about door C2 using Bayes rule:

$$P(C2|H3, G1) = \frac{P(H3|C2, G1)\ P(C2)}{P(H3|G1)} = \frac{1 \times {}^1/_3}{{}^1/_2} = \mathbf{\frac{2}{3}}$$





Digression: Bayes rule - The optimal method of updating knowledge

Case 1: Informed Host: Knows where the car is, and will not open the door that leads to car.

You, the guest, choose door 1. The **Informed** host **consciously** opens door 3 (knowing the car is behind #2).

Prior knowledge about the car being behind door 2 (C2):

$$P(C2|G1) = P(C1|G1) = P(C3|G1) = P(C2) = \frac{1}{3}$$

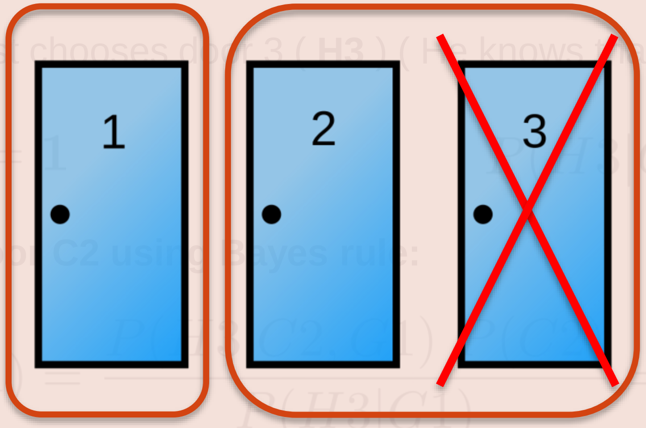
New knowledge: The informed host chooses door 3 (**H3**) (He knows that the car is behind door 2)

$$P(H3|C2,G1) = 1 \qquad P(H3|G1) = \frac{1}{2}$$

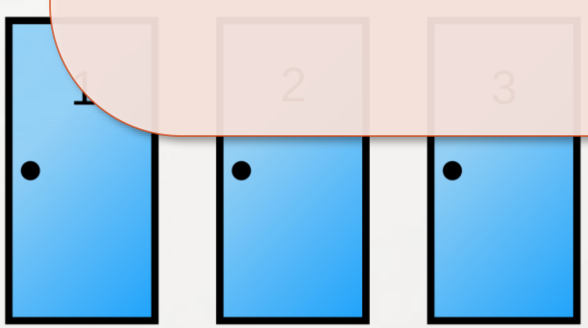
Update your knowledge about door C2 using Bayes rule:

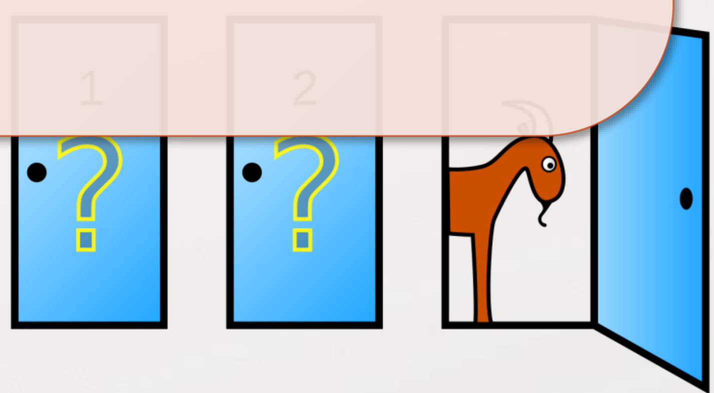
$$P(C2|H3,G1) = \frac{P(H3|C2,G1)\,P(C2)}{P(H3|G1)} = \frac{1 \times 1/3}{1/2} = \frac{2}{3}$$

$$P(C1) = \frac{1}{3}$$



$$P(C2 \cup \cancel{C}3) = \frac{2}{3}$$





Digression: Bayes rule - The optimal method of updating knowledge

Case 1: Uninformed Host

You, the guest, choose door 1. The **Uninformed** host **randomly** opens door 3 (knowing nothing a priori).

Prior knowledge about the car being behind door 2:

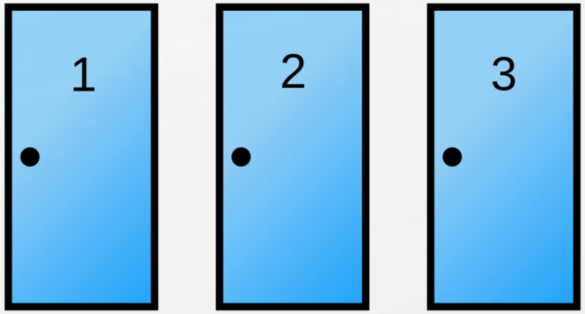
$$P(C2|G1) = P(C1|G1) = P(C3|G1) = P(C2) = \frac{\mathbf{1}}{\mathbf{3}}$$

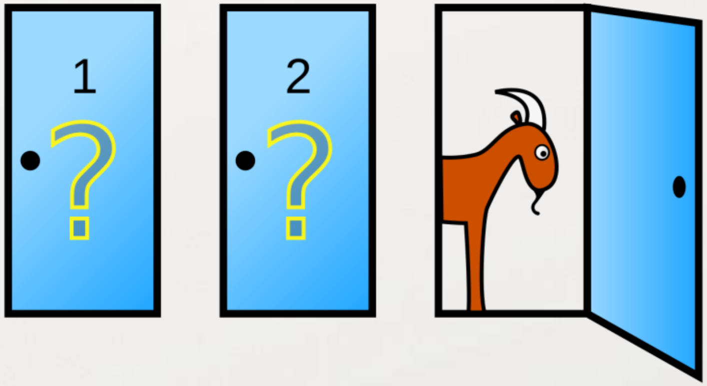
New knowledge: The **Un**informed host chooses door 3 (**H3**) (He does **not** know where the car is)

$$P(H3|C2, G1) = \frac{\mathbf{1}}{\mathbf{2}} \qquad\qquad P(H3|G1) = \frac{\mathbf{1}}{\mathbf{2}}$$

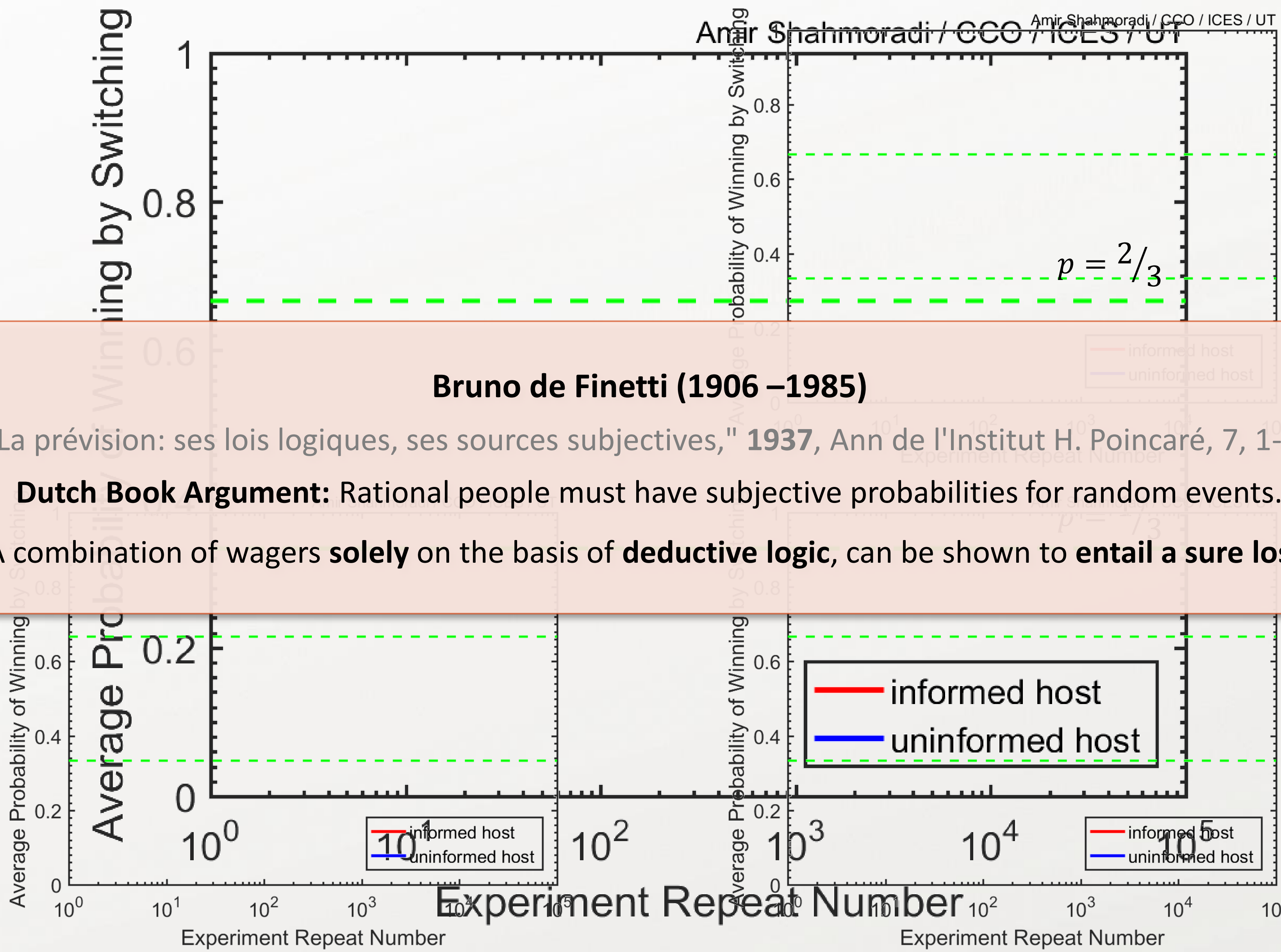
Update your knowledge about door C2 using Bayes rule:

$$P(C2|H3, G1) = \frac{P(H3|C2, G1)\ P(C2)}{P(H3|G1)} = \frac{{}^1\!/\!{}_2 \times {}^1\!/\!{}_3}{{}^1\!/\!{}_2} = \frac{\mathbf{1}}{\mathbf{3}}$$





Digression: Bayes rule - The optimal method of updating knowledge



Bruno de Finetti (1906 –1985)

"La prévision: ses lois logiques, ses sources subjectives," **1937**, Ann de l'Institut H. Poincaré, 7, 1-68.

Dutch Book Argument: Rational people must have subjective probabilities for random events.

A combination of wagers **solely** on the basis of **deductive logic**, can be shown to **entail a sure loss**.

Digression: Bayes rule - The optimal method of updating knowledge

Bayes rule

$$P(B|A,\mathcal{I}) = \frac{P(A|B,\mathcal{I})\ P(B|\mathcal{I})}{P(A|\mathcal{I})}$$

Bayes rule in Bayesian modeling

$$\pi(\boldsymbol{\theta}|\mathcal{D},M) = \frac{\pi(\mathcal{D}|\boldsymbol{\theta},M)\ \pi(\boldsymbol{\theta}|M)}{\pi(\mathcal{D}|M)}$$

Suppose there are multiple rival models,

$$\mathcal{M} = \{M_1, M_2, \ldots, M_K\}$$

Integrate the Bayes rule over the parameter space,

Marginal likelihood (evidence)

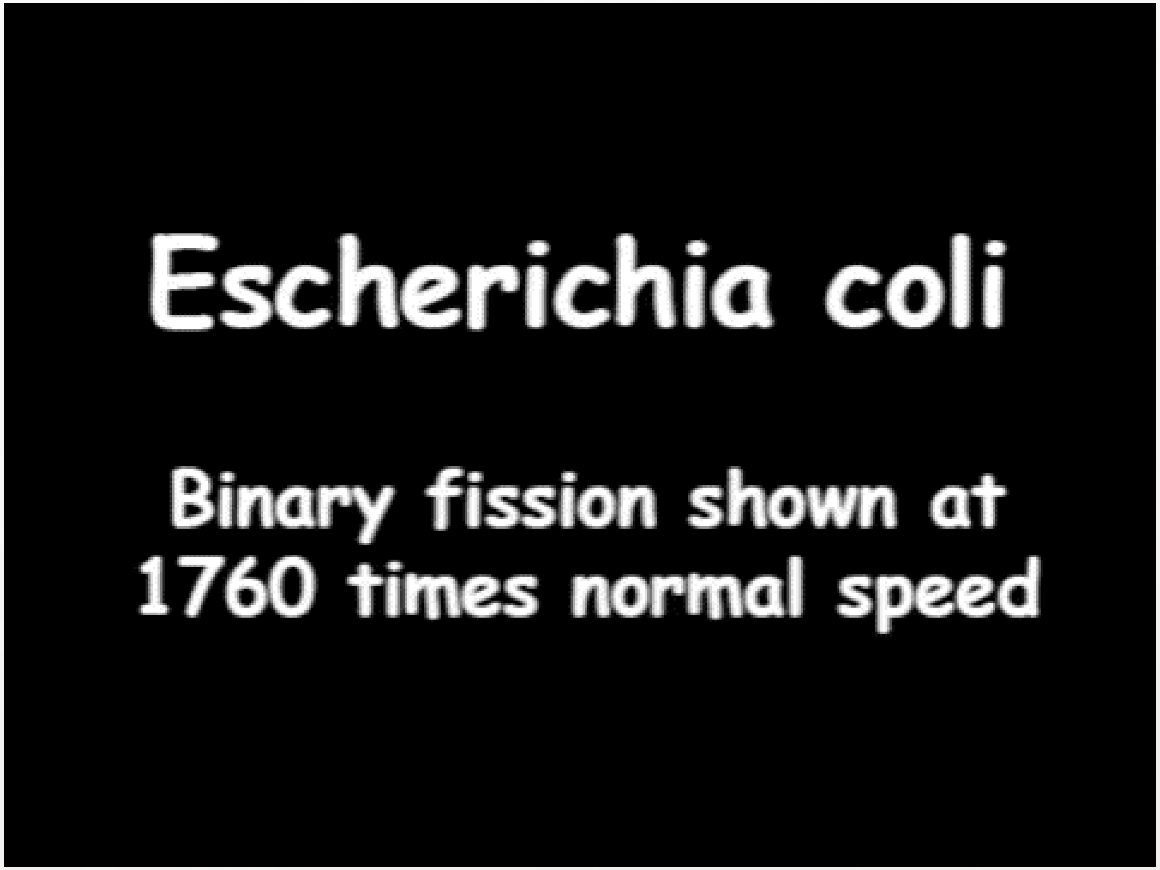
$$1 = \frac{\int_{\boldsymbol{\Theta}} \pi(\mathcal{D}|\boldsymbol{\theta},M_i,\mathcal{M})\ \pi(\boldsymbol{\theta}|M_i,\mathcal{M})\,\mathrm{d}\boldsymbol{\theta}}{\pi(\mathcal{D}|M_i,\mathcal{M})} \Rightarrow \boxed{\pi(D|M_i,\mathcal{M}) = \int_{\boldsymbol{\Theta}} \pi(D|\boldsymbol{\theta},M_i,\mathcal{M})\pi(\boldsymbol{\theta}|M_i,\mathcal{M})\,\mathrm{d}\boldsymbol{\theta}}$$

Apply the Bayes rule, this time, on the discrete set of models $\mathcal{M}$,

$$\pi(M_i|\mathcal{D},\mathcal{M}) = \frac{\pi(\mathcal{D}|M_i)\ \pi(M_i|\mathcal{M})}{\pi(\mathcal{D}|\mathcal{M})}$$

All models are wrong, but some are useful

Example problem: Modeling the growth of cells and bacteria

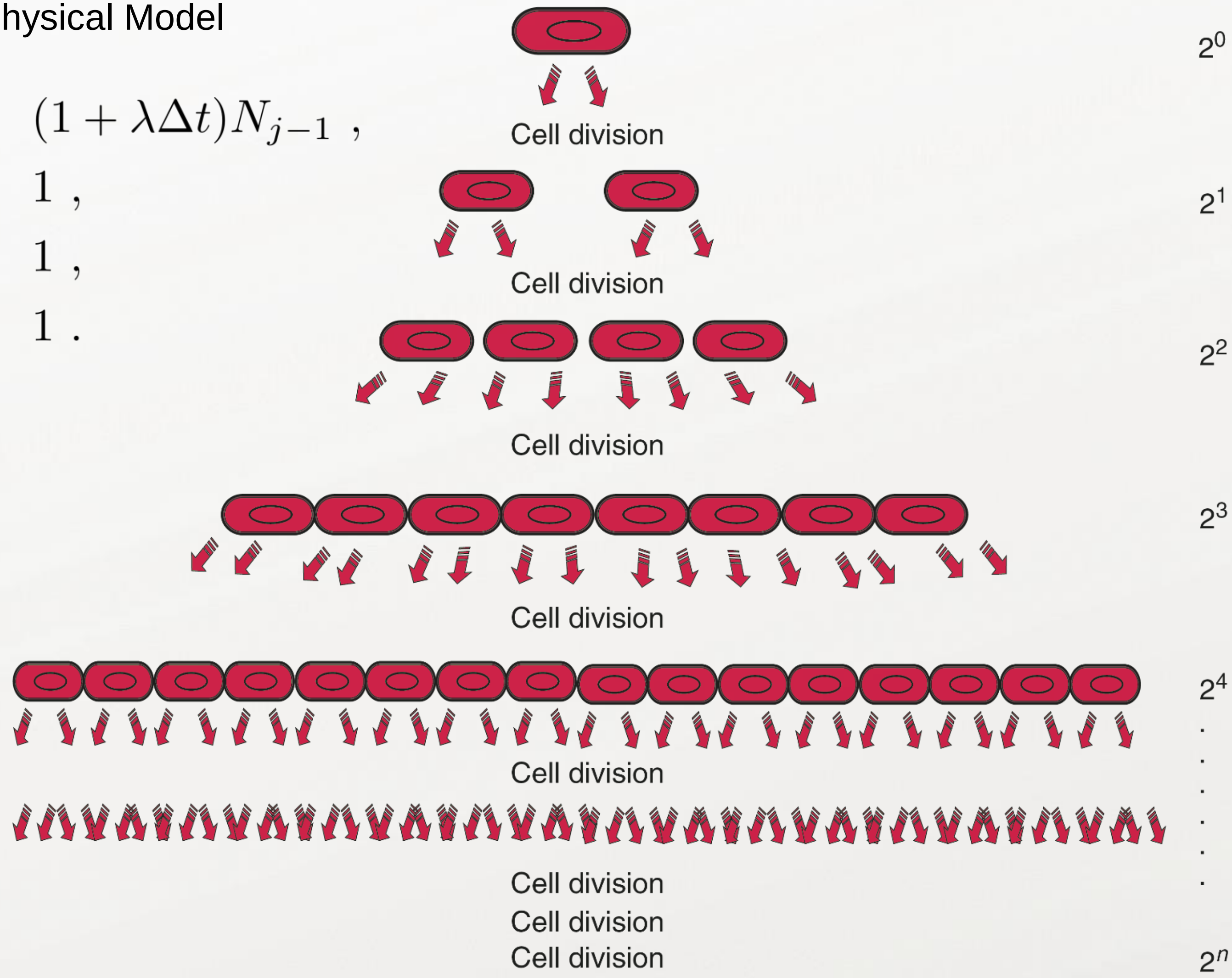


All models are wrong, but some are useful

Cell division is a multiplicative process.

The Physical Model

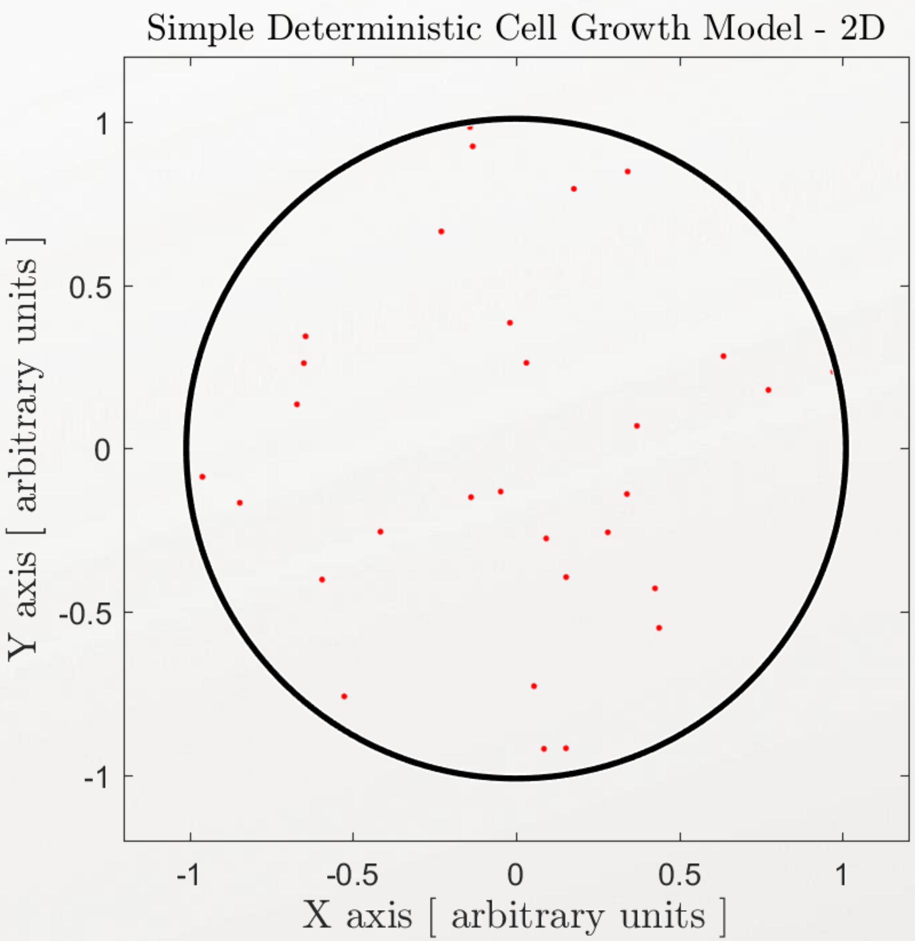
$$
\begin{aligned}
N_j &= (1+\lambda\Delta t)N_{j-1}\ , \\
N_0 &= 1\ , \\
\lambda &= 1\ , \\
\Delta t &= 1\ .
\end{aligned}
$$

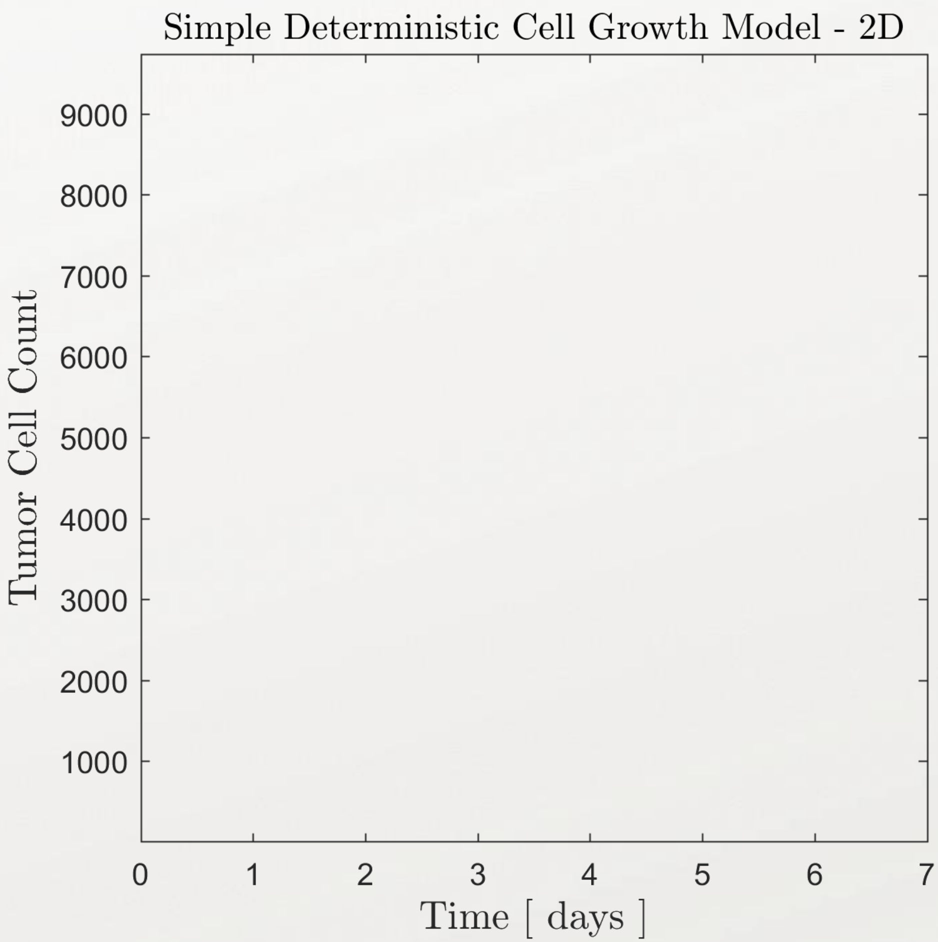


All models are wrong, but some are useful

For the moment, suppose we live in a deterministic ideal world, with,
1. observational data that contain **no uncertainty** whatsoever.
2. **ideal models** that perfectly match observational data

$$\mathrm{d}\,N = \lambda N \,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{\lambda t}$$



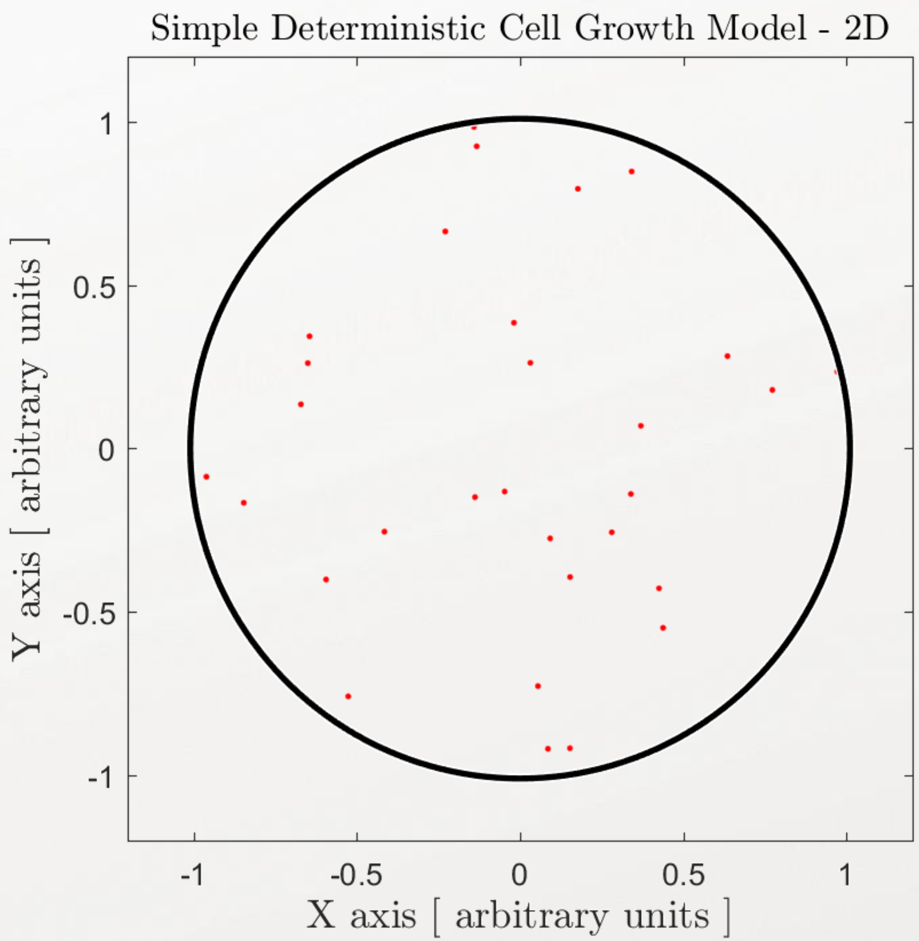


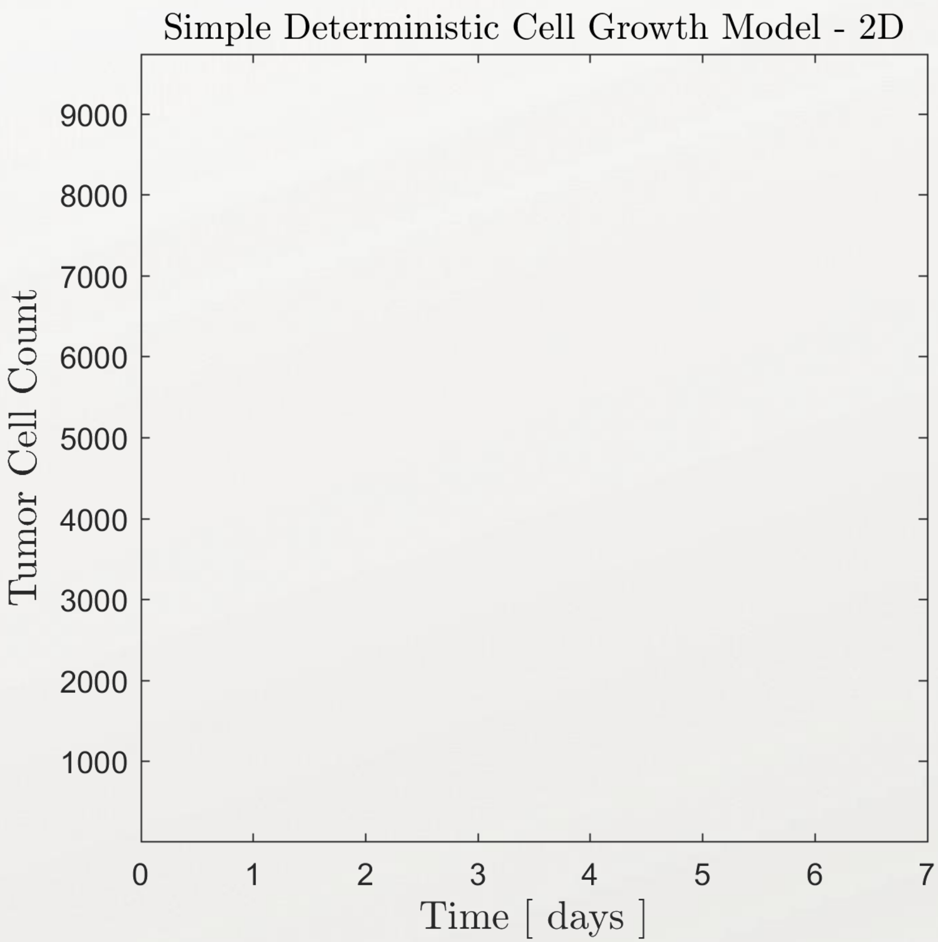
All models are wrong, but some are useful

For the moment, suppose we live in a deterministic ideal world, with,

1. observational data that contain **no uncertainty** whatsoever.
2. **ideal models** that perfectly match observational data

$$\mathrm{d}\,N = \lambda N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{\lambda t}$$





All models are wrong, but some are useful

For the moment, suppose we live in a deterministic ideal world, with,
1. observational data that contain **no uncertainty** whatsoever.
2. **ideal models** that perfectly match observational data

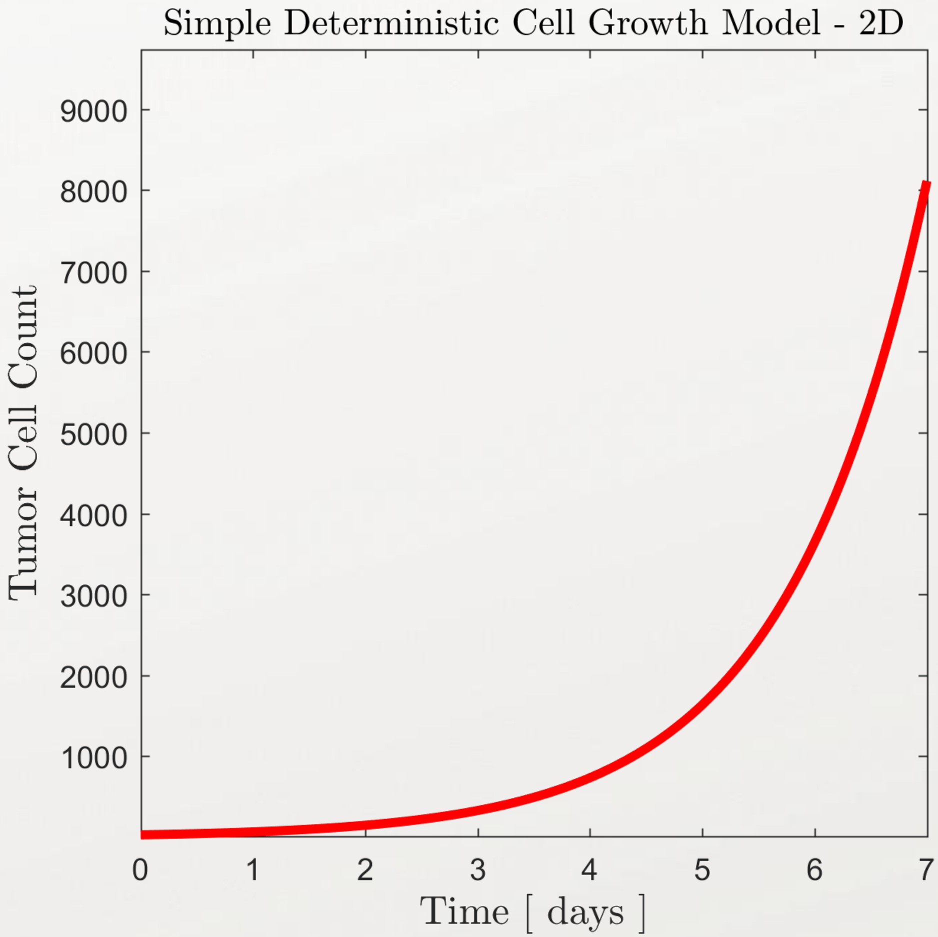
$$\mathrm{d}\,N = \lambda N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{\lambda t}$$

$n_{\mathrm{dv}} = 5$ variables

$\mathcal{R}$	$\boldsymbol{X}$	$\boldsymbol{Y}$	$\boldsymbol{Z}$	$\boldsymbol{t}$	$\boldsymbol{N}$
$\boldsymbol{R}_1$	x_1	y_1	z_1	t_1	N_1
$\boldsymbol{R}_2$	x_2	y_2	z_2	t_2	N_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{R}_{n_{\mathrm{do}}}$	$x_{n_{\mathrm{do}}}$	$y_{n_{\mathrm{do}}}$	$z_{n_{\mathrm{do}}}$	$t_{n_{\mathrm{do}}}$	$N_{n_{\mathrm{do}}}$

$$\begin{aligned}\mathcal{R} &= \{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_{\mathrm{n_{do}}}\}\ ,\\ \boldsymbol{R}_{\mathrm{i}} &= \left\{\boldsymbol{R}_{\mathrm{i,ind}}, \boldsymbol{R}_{\mathrm{i,dep}}\left(\boldsymbol{R}_{\mathrm{i,ind}}\right)\right\}\ .\end{aligned}$$

$$\boxed{\begin{aligned}\boldsymbol{R}_{\mathrm{i}} &- \boldsymbol{M}_{\mathrm{phys}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}\right)\\ &= \boldsymbol{0} \quad \forall\ \boldsymbol{R}_{\mathrm{i}} \in \mathcal{R}\ .\end{aligned}}$$



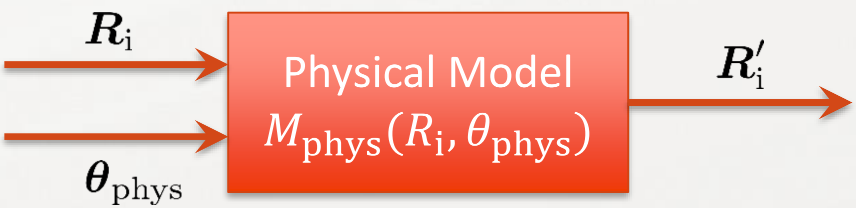
All models are wrong, but some are useful

For the moment, suppose we live in a deterministic ideal world, with,
1. observational data that contain **no uncertainty** whatsoever.
2. **ideal models** that perfectly match observational data

$$\mathrm{d}\,N = \lambda N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{\lambda t}$$

$n_{\mathrm{dv}} = 5$ variables

$\mathcal{R}$	$\boldsymbol{X}$	$\boldsymbol{Y}$	$\boldsymbol{Z}$	$\boldsymbol{t}$	$\boldsymbol{N}$
$\boldsymbol{R}_1$	x_1	y_1	z_1	t_1	N_1
$\boldsymbol{R}_2$	x_2	y_2	z_2	t_2	N_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{R}_{n_{\mathrm{do}}}$	$x_{n_{\mathrm{do}}}$	$y_{n_{\mathrm{do}}}$	$z_{n_{\mathrm{do}}}$	$t_{n_{\mathrm{do}}}$	$N_{n_{\mathrm{do}}}$



$$\begin{aligned}
\mathcal{R} &= \{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_{n_{\mathrm{do}}}\}\ , \\
\boldsymbol{R}_{\mathrm{i}} &= \{\boldsymbol{R}_{\mathrm{i,ind}}, \boldsymbol{R}_{\mathrm{i,dep}}(\boldsymbol{R}_{\mathrm{i,ind}})\}\ .
\end{aligned}$$

$$\boxed{\begin{aligned}
\boldsymbol{R}_{\mathrm{i}} &- \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}) \\
&= \boldsymbol{0} \quad \forall\ \boldsymbol{R}_{\mathrm{i}} \in \mathcal{R}\ .
\end{aligned}}$$

All models are wrong, but some are useful

Cell division is a multiplicative process.

The Physical Model

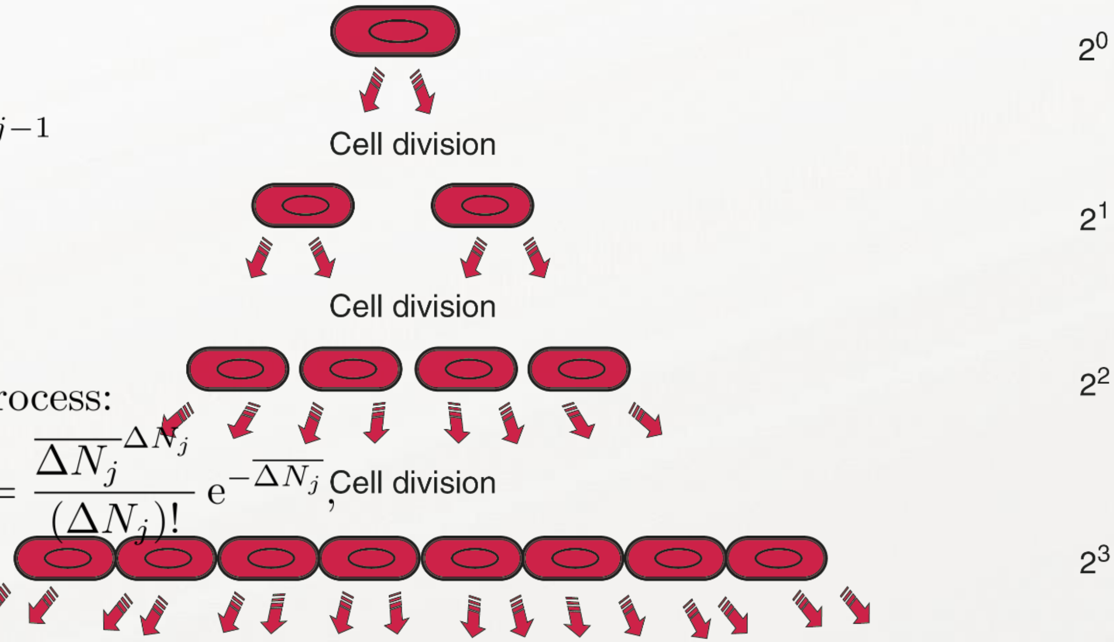
$$
\begin{aligned}
N_j &= (1+\lambda\Delta t)N_{j-1}\\
N_0 &= 1\\
\lambda &= \text{stochatic}\\
&\quad\ \text{variable}
\end{aligned}
$$

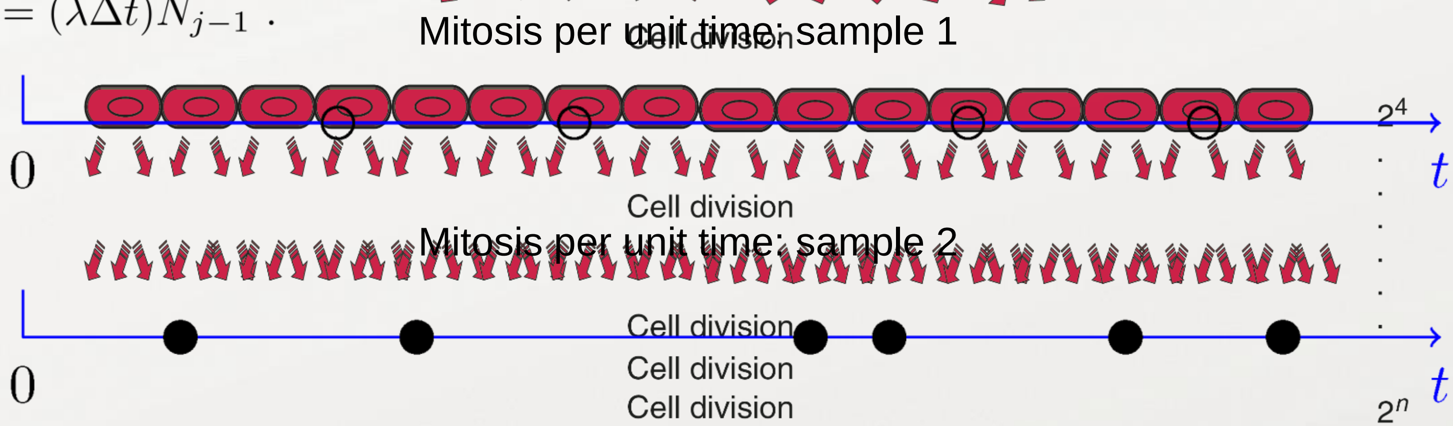
Poisson Stochastic Process:

$$\pi(\Delta N_j | N_{j-1}, \overline{\lambda}, \Delta t) = \frac{\overline{\Delta N_j}^{\Delta N_j}}{(\Delta N_j)!}\, \mathrm{e}^{-\overline{\Delta N_j}},$$

$$\Delta N_j = (\lambda \Delta t) N_{j-1}\ ,$$

$$\overline{\Delta N_j} = (\overline{\lambda}\Delta t) N_{j-1}\ .$$





All models are wrong, but some are useful

Cell division is a multiplicative process.

The Physical Model

$$
\begin{aligned}
N_j &= (1+\lambda\Delta t)N_{j-1} \\
N_0 &= 1 \\
\lambda &= \text{stochastic} \\
&\quad\ \text{variable}
\end{aligned}
$$

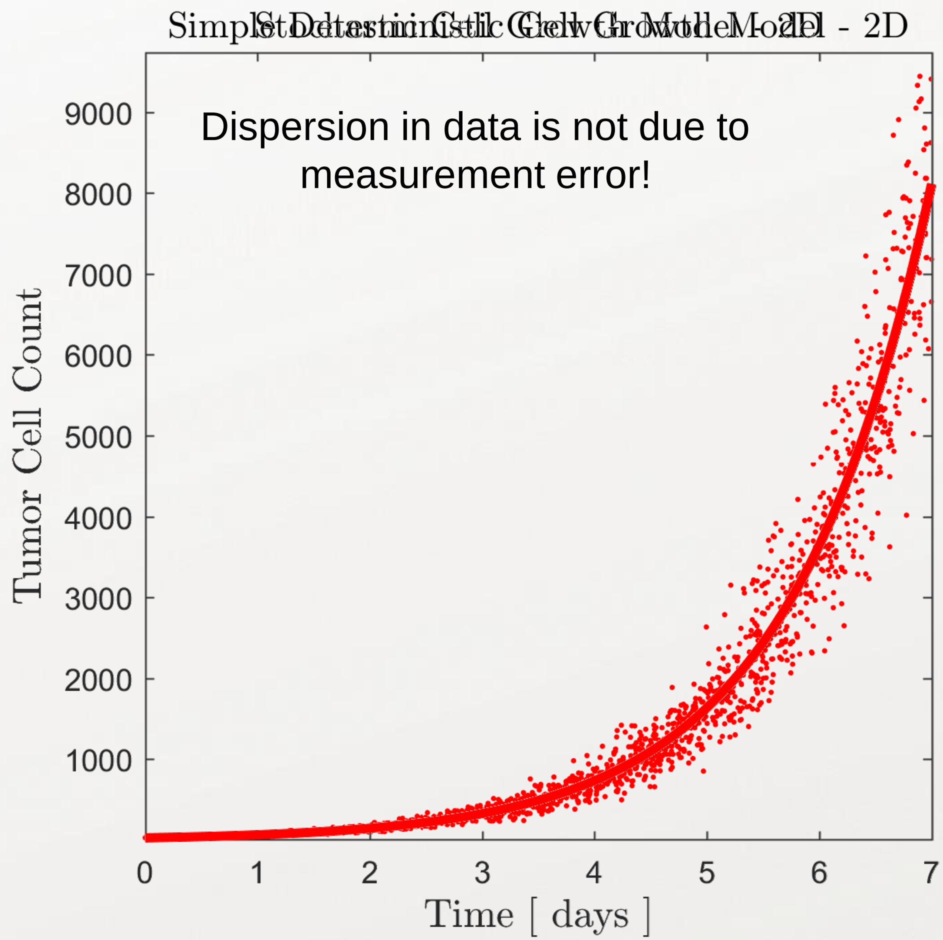
Poisson Stochastic Process:

$$\pi(\Delta N_j | N_{j-1}, \overline{\lambda}, \Delta t) = \frac{\overline{\Delta N_j}^{\Delta N_j}}{(\Delta N_j)!}\, \mathrm{e}^{-\overline{\Delta N_j}},$$

$$\Delta N_j = (\lambda \Delta t) N_{j-1} \ ,$$

$$\overline{\Delta N_j} = (\overline{\lambda}\Delta t) N_{j-1} \ .$$

$$
\boxed{
\begin{aligned}
\boldsymbol{R}_\mathrm{i} \quad &- \quad \boldsymbol{M}_\mathrm{phys}(\boldsymbol{R}_\mathrm{i}, \boldsymbol{\theta}_\mathrm{phys}, \mathcal{S}_\mathrm{phys}) \\
&\neq \quad \mathbf{0} \quad \forall\ \boldsymbol{R}_\mathrm{i} \in \mathcal{R} \ .
\end{aligned}
}
$$



All models are wrong, but some are useful

What is the useful model?

The Physical Model

$$
\begin{aligned}
N_j &= (1+\lambda\Delta t)N_{j-1} \\
N_0 &= 1 \\
\lambda &= \text{stochastic} \\
&\quad\ \text{variable}
\end{aligned}
$$

Poisson Stochastic Process:

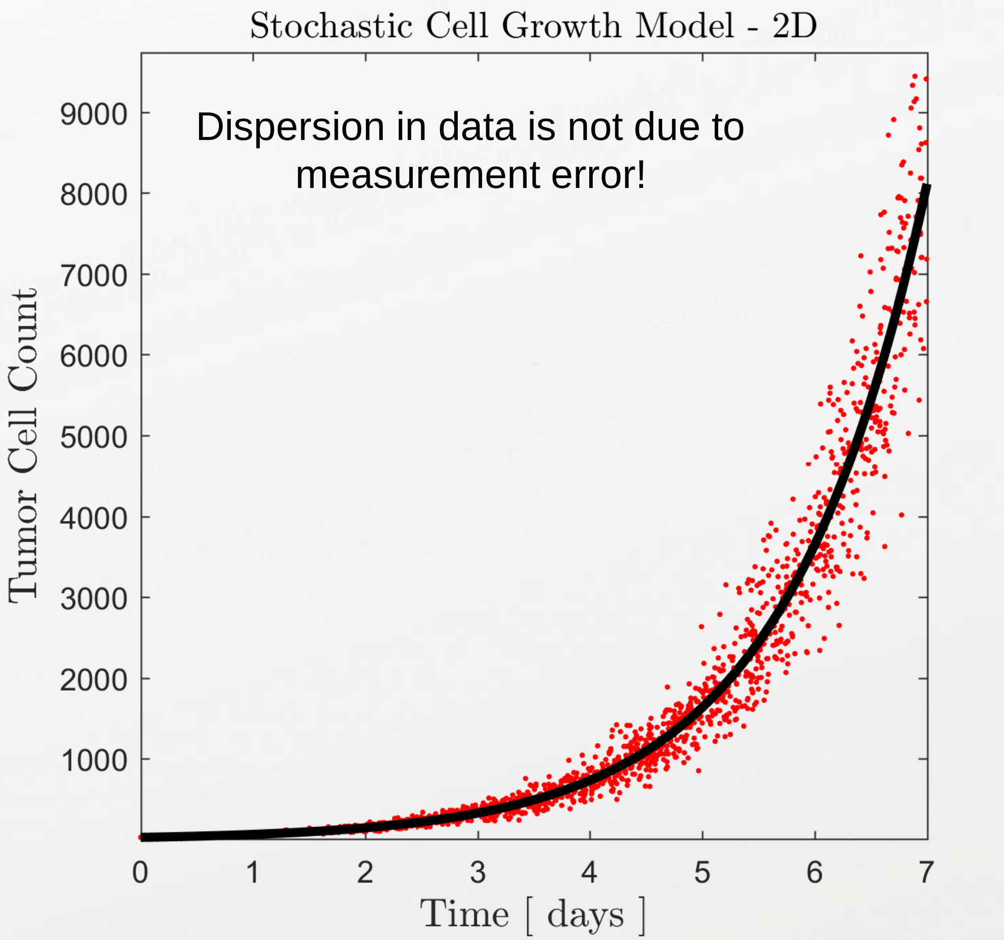
$$\pi(\Delta N_j | N_{j-1}, \overline{\lambda}, \Delta t) = \frac{\overline{\Delta N_j}^{\Delta N_j}}{(\Delta N_j)!}\, \mathrm{e}^{-\overline{\Delta N_j}},$$

$$\Delta N_j = (\lambda \Delta t) N_{j-1}\ ,$$

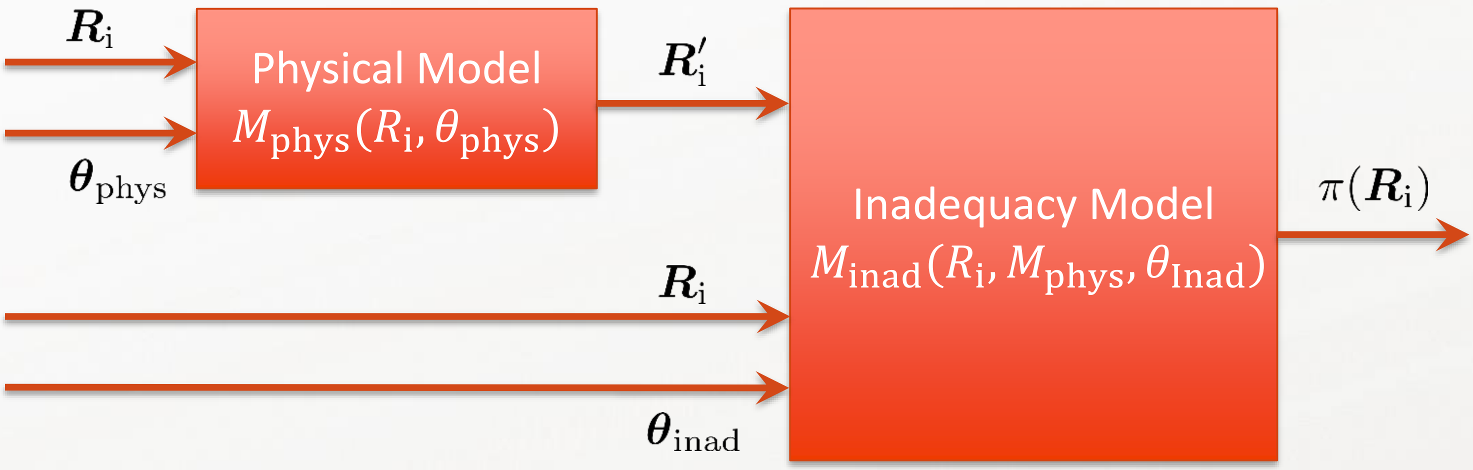
$$\overline{\Delta N_j} = (\overline{\lambda}\Delta t) N_{j-1}\ .$$

$$\boxed{\begin{aligned} \boldsymbol{R}_{\mathrm{i}} &- \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}) \\ &\neq \mathbf{0} \quad \forall\ \boldsymbol{R}_{\mathrm{i}} \in \mathcal{R}\ . \end{aligned}}$$

$$\mathrm{d}\,N = \overline{\lambda} N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{\overline{\lambda} t}$$



All physical models need stochastic *inadequacy models*



$$\mathcal{R} = \{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{\rm n_{do}}\}$$

$\mathcal{R}$	$\boldsymbol{X}$	$\boldsymbol{Y}$	$\boldsymbol{Z}$	$\boldsymbol{t}$	$\boldsymbol{N}$
$\boldsymbol{R}_1$	x_1	y_1	z_1	t_1	N_1
$\boldsymbol{R}_2$	x_2	y_2	z_2	t_2	N_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{R}_{n_{do}}$	$x_{n_{do}}$	$y_{n_{do}}$	$z_{n_{do}}$	$t_{n_{do}}$	$N_{n_{do}}$

$$\begin{aligned}
\boldsymbol{\theta}_{\rm pi} &= \{\boldsymbol{\theta}_{\rm phys}, \boldsymbol{\theta}_{\rm inad}\} \ , \\
\boldsymbol{M}_{\rm pi} &= \{\boldsymbol{M}_{\rm phys}, \boldsymbol{M}_{\rm inad}\} \ , \\
\pi(\mathcal{R}|\boldsymbol{\theta}_{\rm pi}, \boldsymbol{M}_{\rm pi}) &= \prod_{\rm i=1}^{n_{\rm do}} \pi(\boldsymbol{R}_{\rm i}) \ .
\end{aligned}$$

All physical models need stochastic *inadequacy models*

Least-Squares regression is the most abused statistical method in history.
Normal distribution is the most abused statistical distribution in history.
Normal likelihood is the most abused statistical distribution in history.

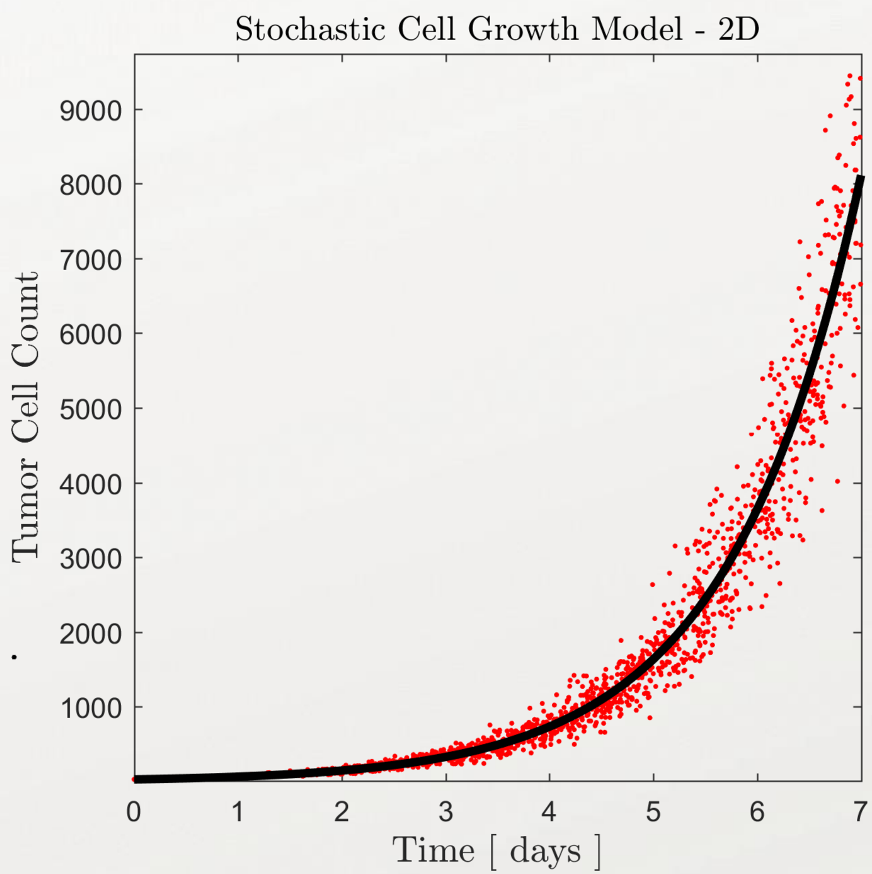
$$\boldsymbol{R}_{\mathrm{i}} - \boldsymbol{M}_{\mathrm{phys}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}\right) = \boldsymbol{U}_{\mathrm{i}} ~,$$

$$\begin{aligned} \boldsymbol{U}_{\mathrm{i}} &\sim \boldsymbol{M}_{\mathrm{inad}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{inad}}, \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}})\right) \\ &\sim \mathcal{N}\left(\mu = \boldsymbol{M}_{\mathrm{phys}}, \sigma\right) ~, \end{aligned}$$

$$\begin{aligned} \mathcal{U} &= \{\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{\mathrm{n_{do}}}\} ~, \\ \boldsymbol{M}_{\mathrm{pi}} &= \{\boldsymbol{M}_{\mathrm{phys}}, \boldsymbol{M}_{\mathrm{inad}}\} ~, \\ \boldsymbol{\theta}_{\mathrm{pi}} &= \{\boldsymbol{\theta}_{\mathrm{phys}}, \boldsymbol{\theta}_{\mathrm{inad}}\} ~, \end{aligned}$$

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) &\equiv \pi\left(\mathcal{U} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \\ &\overset{\mathrm{i.i.d.}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \pi\left(\boldsymbol{U}_{\mathrm{i}} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) ~, \\ &\overset{\mathcal{N}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\big[\overbrace{N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})}^{\boldsymbol{U}_{\mathrm{i}}}\big]^2}{2\sigma^2} \right) ~, \end{aligned}$$

$$\Rightarrow \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) = \underbrace{- \sum_{\mathrm{i}=1}^{n_{\mathrm{do}}} \left[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\right]^2}_{\text{least−squares cost function}} + \cdots ~.$$



All physical models need stochastic *inadequacy models*

Least-Squares regression is the most abused statistical method in history.
Normal distribution is the most abused statistical distribution in history.
Normal likelihood is the most abused statistical distribution in history.

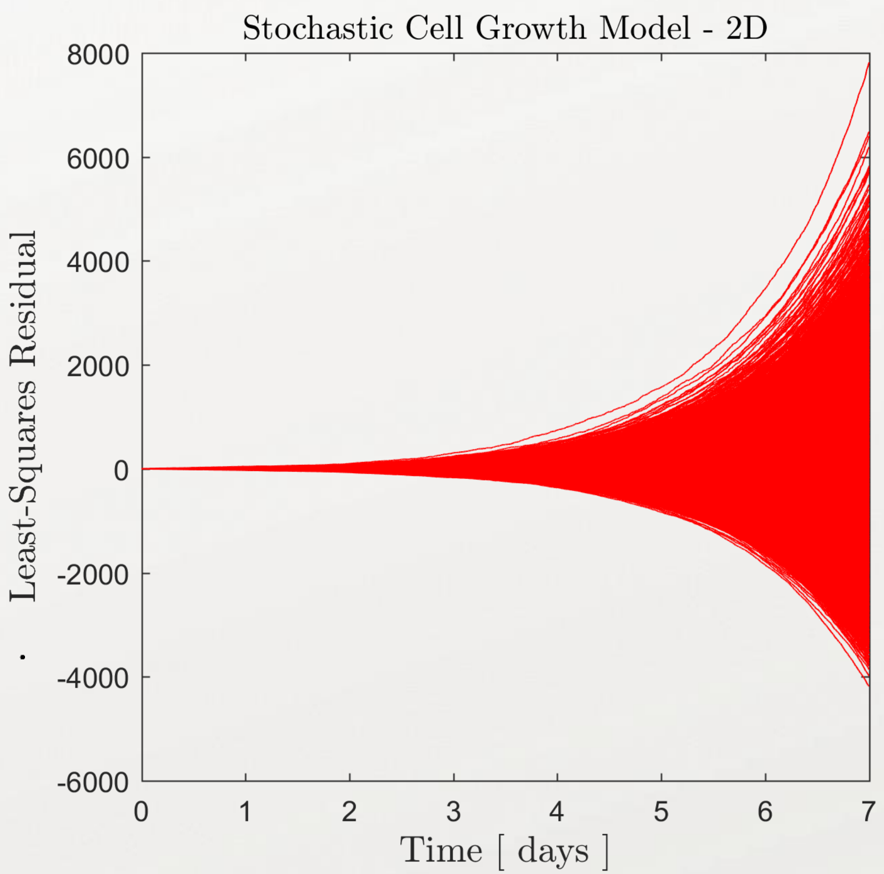
$$\boldsymbol{R}_{\mathrm{i}} - \boldsymbol{M}_{\mathrm{phys}}\big(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}\big) = \boldsymbol{U}_{\mathrm{i}} \ , \qquad \boldsymbol{U}_{\mathrm{i}} \sim \boldsymbol{M}_{\mathrm{inad}}\big(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{inad}}, \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}})\big)$$

$$\sim \mathcal{N}\big(\mu = \boldsymbol{M}_{\mathrm{phys}}, \sigma\big) \ ,$$

$$\begin{aligned} \mathcal{U} &= \{\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{\mathrm{n_{do}}}\} \ , \\ \boldsymbol{M}_{\mathrm{pi}} &= \{\boldsymbol{M}_{\mathrm{phys}}, \boldsymbol{M}_{\mathrm{inad}}\} \ , \\ \boldsymbol{\theta}_{\mathrm{pi}} &= \{\boldsymbol{\theta}_{\mathrm{phys}}, \boldsymbol{\theta}_{\mathrm{inad}}\} \ , \end{aligned}$$

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}};\mathcal{U}) &\equiv \pi\big(\mathcal{U}\big|\boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\big) \\ &\overset{\mathrm{i.i.d.}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \pi\big(\boldsymbol{U}_{\mathrm{i}}\big|\boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\big) \ , \\ &\overset{\mathcal{N}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\overbrace{\big[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\big]}^{\boldsymbol{U}_{\mathrm{i}}}{}^2}{2\sigma^2}\right) \ , \end{aligned}$$

$$\Rightarrow \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}};\mathcal{U}) = \underbrace{-\sum_{\mathrm{i}=1}^{n_{\mathrm{do}}} \big[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\big]^2}_{\text{least-squares cost function}} + \cdots \ .$$



All physical models need stochastic *inadequacy models*

Least-Squares regression is the most abused statistical method in history.
Normal distribution is the most abused statistical distribution in history.
Normal likelihood is the most abused statistical distribution in history.

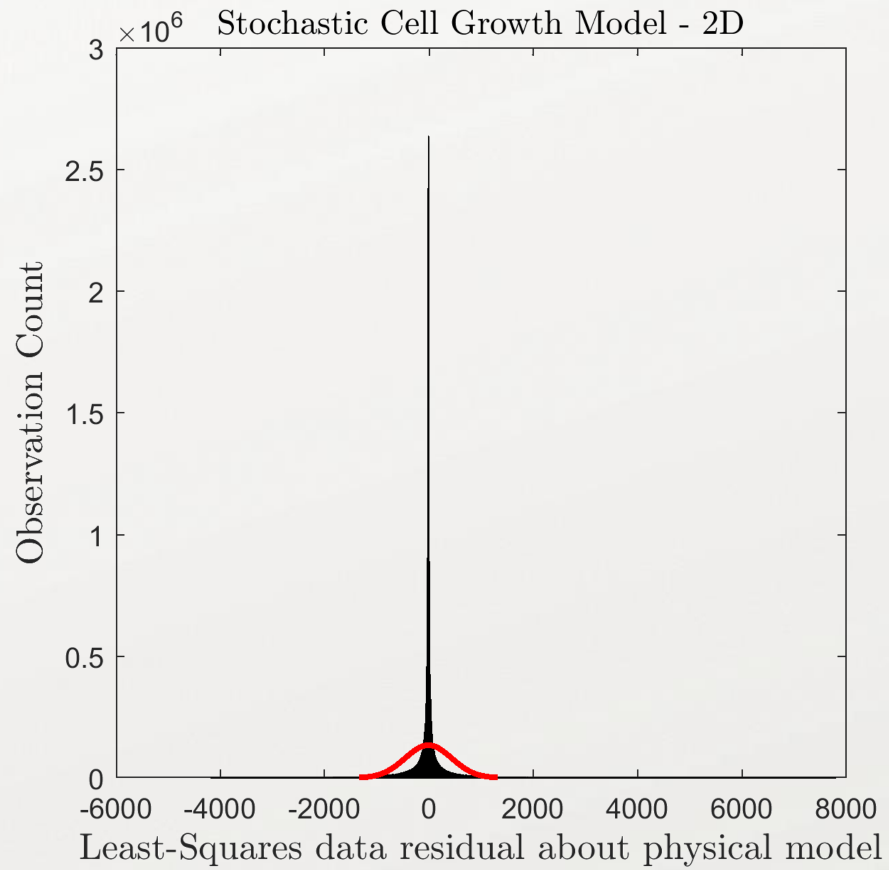
$$\boldsymbol{R}_{\mathrm{i}} - \boldsymbol{M}_{\mathrm{phys}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}\right) = \boldsymbol{U}_{\mathrm{i}} \ ,$$

$$\begin{aligned} \boldsymbol{U}_{\mathrm{i}} &\sim \boldsymbol{M}_{\mathrm{inad}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{inad}}, \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}})\right) \\ &\sim \mathcal{N}\left(\mu = \boldsymbol{M}_{\mathrm{phys}}, \sigma\right) \ , \end{aligned}$$

$$\begin{aligned} \mathcal{U} &= \left\{\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{\mathrm{n_{do}}}\right\} \ , \\ \boldsymbol{M}_{\mathrm{pi}} &= \left\{\boldsymbol{M}_{\mathrm{phys}}, \boldsymbol{M}_{\mathrm{inad}}\right\} \ , \\ \boldsymbol{\theta}_{\mathrm{pi}} &= \left\{\boldsymbol{\theta}_{\mathrm{phys}}, \boldsymbol{\theta}_{\mathrm{inad}}\right\} \ , \end{aligned}$$

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) &\equiv \pi\left(\mathcal{U} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \\ &\overset{\mathrm{i.i.d.}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \pi\left(\boldsymbol{U}_{\mathrm{i}} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \ , \\ &\overset{\mathcal{N}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\overbrace{\left[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\right]}^{\boldsymbol{U}_{\mathrm{i}}}{}^2}{2\sigma^2} \right) \ , \end{aligned}$$

$$\Rightarrow \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) = \underbrace{- \sum_{\mathrm{i}=1}^{n_{\mathrm{do}}} \left[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\right]^2}_{\text{least-squares cost function}} + \cdots \ .$$



All physical models need stochastic *inadequacy models*

Least-Squares regression is the most abused statistical method in history.
Normal distribution is the most abused statistical distribution in history.
Normal likelihood is the most abused statistical distribution in history.

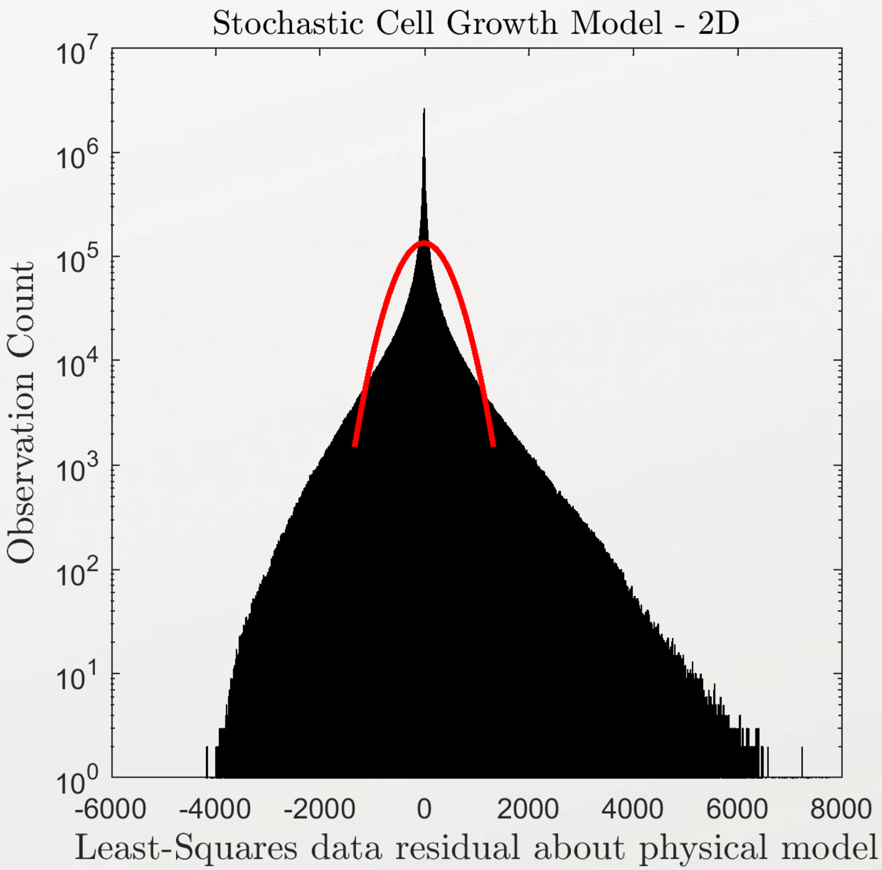
$$\boldsymbol{R}_{\mathrm{i}} - \boldsymbol{M}_{\mathrm{phys}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}}\right) = \boldsymbol{U}_{\mathrm{i}} ~,$$

$$\begin{aligned} \boldsymbol{U}_{\mathrm{i}} &\sim \boldsymbol{M}_{\mathrm{inad}}\left(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{inad}}, \boldsymbol{M}_{\mathrm{phys}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{phys}}, \mathcal{S}_{\mathrm{phys}})\right) \\ &\sim \mathcal{N}\left(\mu = \boldsymbol{M}_{\mathrm{phys}}, \sigma\right) ~, \end{aligned}$$

$$\begin{aligned} \mathcal{U} &= \left\{\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{\mathrm{n_{do}}}\right\} ~, \\ \boldsymbol{M}_{\mathrm{pi}} &= \left\{\boldsymbol{M}_{\mathrm{phys}}, \boldsymbol{M}_{\mathrm{inad}}\right\} ~, \\ \boldsymbol{\theta}_{\mathrm{pi}} &= \left\{\boldsymbol{\theta}_{\mathrm{phys}}, \boldsymbol{\theta}_{\mathrm{inad}}\right\} ~, \end{aligned}$$

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) &\equiv \pi\left(\mathcal{U} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \\ &\overset{\mathrm{i.i.d.}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \pi\left(\boldsymbol{U}_{\mathrm{i}} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) ~, \\ &\overset{\mathcal{N}}{=} \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\overbrace{\left[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\right]}^{\boldsymbol{U}_{\mathrm{i}}}{}^2}{2\sigma^2} \right) ~, \end{aligned}$$

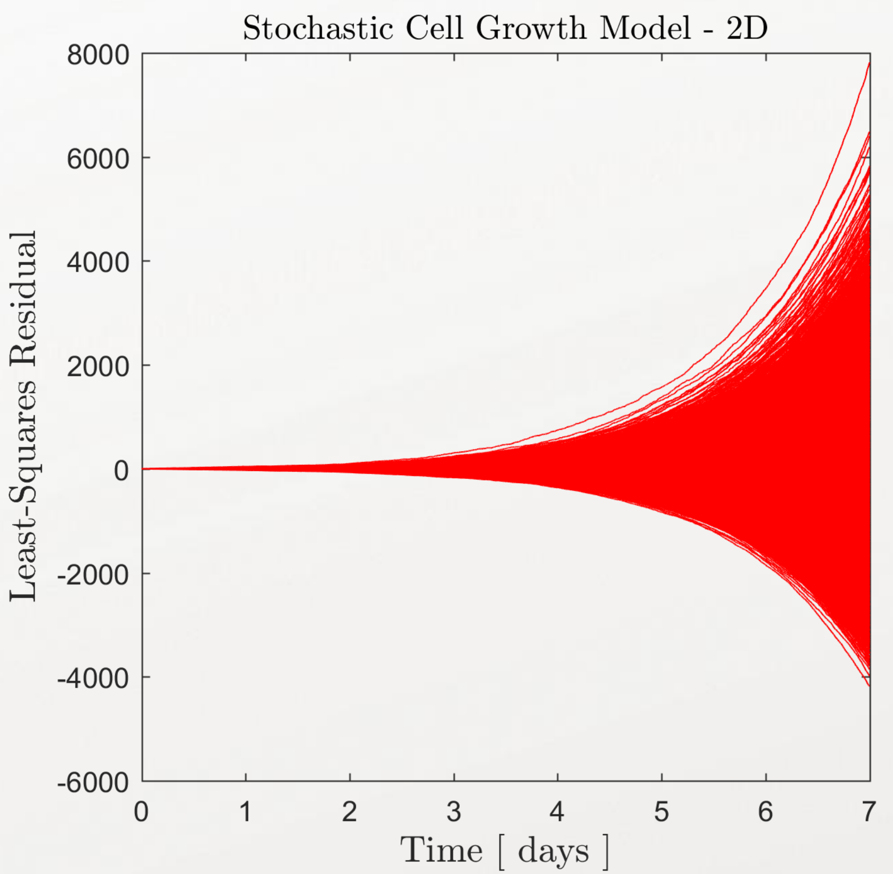
$$\Rightarrow \mathcal{L}(\boldsymbol{\theta}_{\mathrm{pi}}; \mathcal{U}) = \underbrace{- \sum_{\mathrm{i}=1}^{n_{\mathrm{do}}} \left[N(t_{\mathrm{i}}) - \boldsymbol{M}_{\mathrm{phys}}(t_{\mathrm{i}})\right]^2}_{\text{least−squares cost function}} + \cdots ~.$$

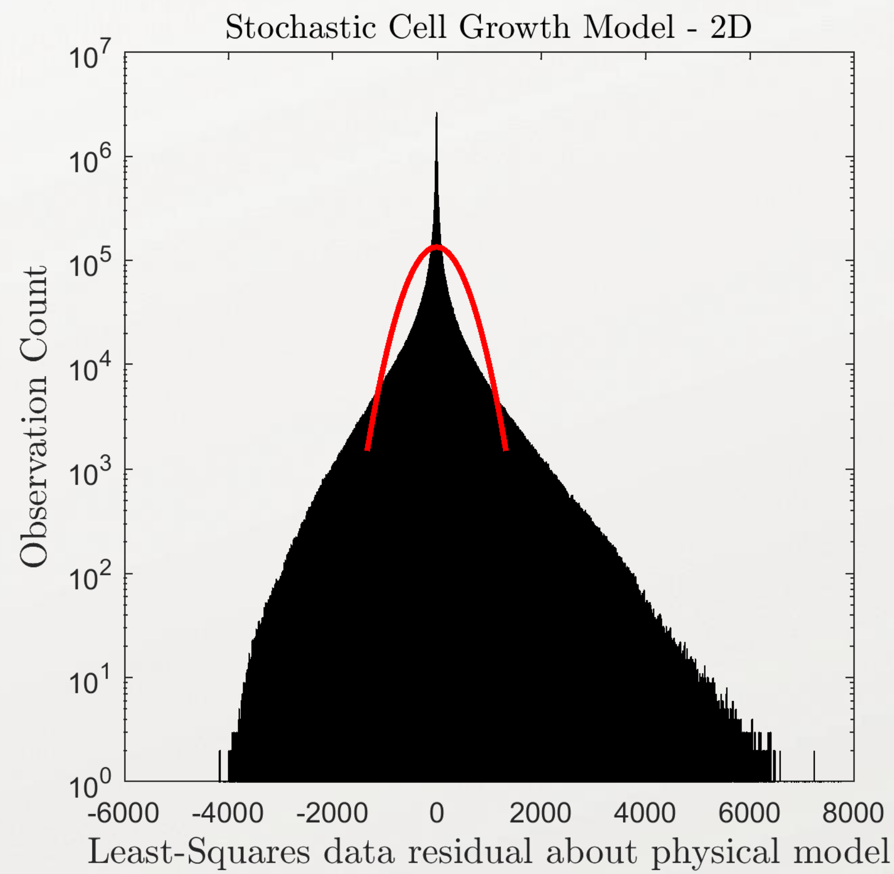


Common flaws associated with Least-Squares approach in tumor modeling

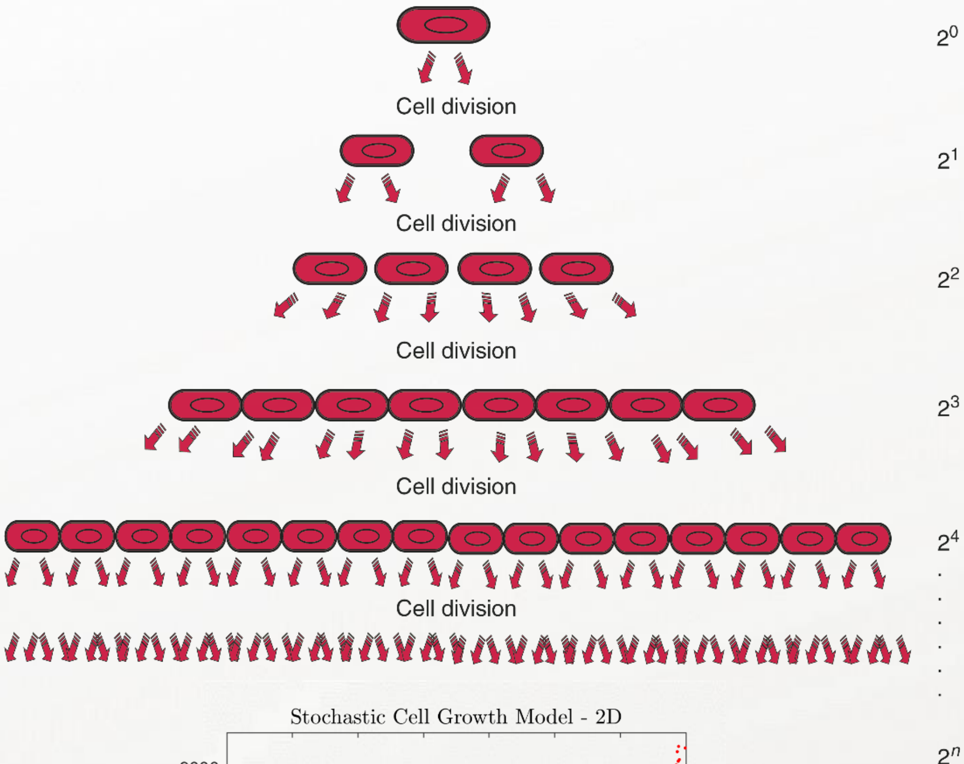
Least-Squares approach in tumor modeling leads to:

1. **False-positive** predictions for the future growth of tumor (biased toward high tumor cell count)
2. **False-negative** predictions for the future decline of tumor (biased toward low tumor cell count)
3. **Overestimation** of prediction **uncertainty** early-on (negative initial cell count!)
4. **Underestimation** of prediction **uncertainty** in distant future.

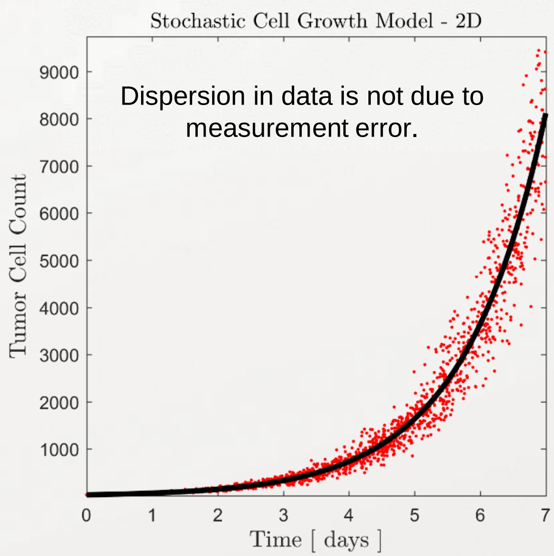




Tumor growth modeling requires inadequacy models other than Least-Squares



$$
\begin{aligned}
N_j &= \lambda_j N_{j-1} \\
&= \lambda_j \big(\lambda_{j-1} N_{j-2}\big) \\
&= \lambda_j \big(\lambda_{j-1} \big(\lambda_{j-2} N_{j-3}\big)\big) \\
&\ \vdots \\
&= N_0 \times \prod_{i=1}^{j} \lambda_j \\
\Rightarrow \ln N_j &= \ln N_0 + \underbrace{\sum_{i=1}^{j} \ln \lambda_i}_{\textit{sum of independent random variables}}
\end{aligned}
$$



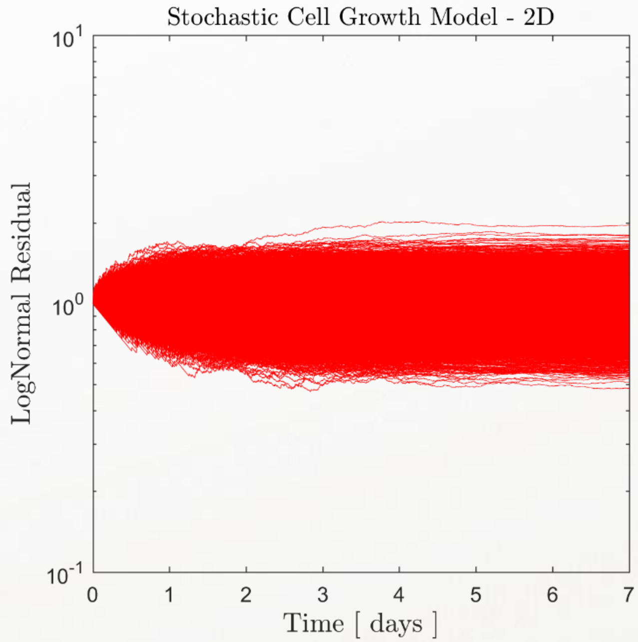
Central Limit Theorem

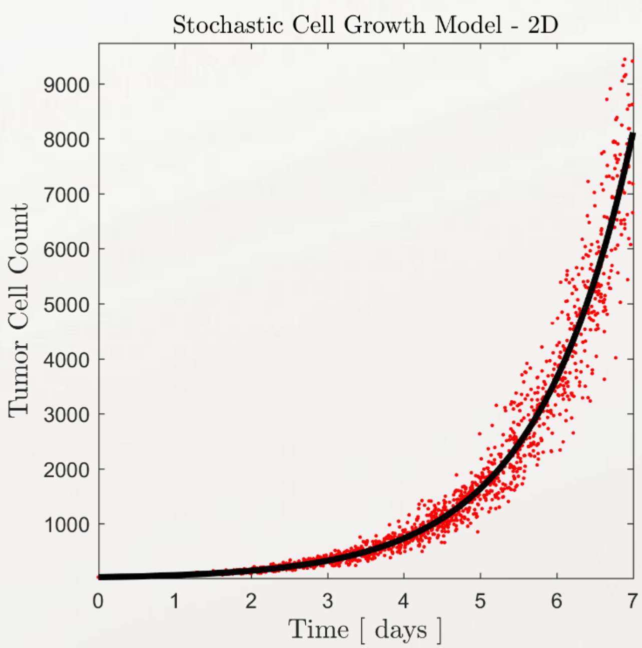
Sum of n independent random variables tends to normal distribution as $n \to \infty$.

$$
\Rightarrow \sum_{i=1}^{j} \ln \lambda_i \sim \mathcal{N}\big(\overline{\ln \lambda}, \sigma\big) \ ,
$$

$$
\Rightarrow \prod_{i=1}^{j} \lambda_i \sim \mathcal{LN}\big(\overline{\ln \lambda}, \sigma\big) \ .
$$

Tumor growth modeling requires inadequacy models other than Least-Squares





$$
\begin{aligned}
N_j &= \lambda_j N_{j-1} \\
&= \lambda_j \big(\lambda_{j-1} N_{j-2}\big) \\
&= \lambda_j \big(\lambda_{j-1} \big(\lambda_{j-2} N_{j-3}\big)\big) \\
&\ \vdots \\
&= N_0 \times \prod_{i=1}^{j} \lambda_j \\
\Rightarrow \ln N_j &= \ln N_0 + \underbrace{\sum_{i=1}^{j} \ln \lambda_i}_{\textit{sum of independent random variables}}
\end{aligned}
$$

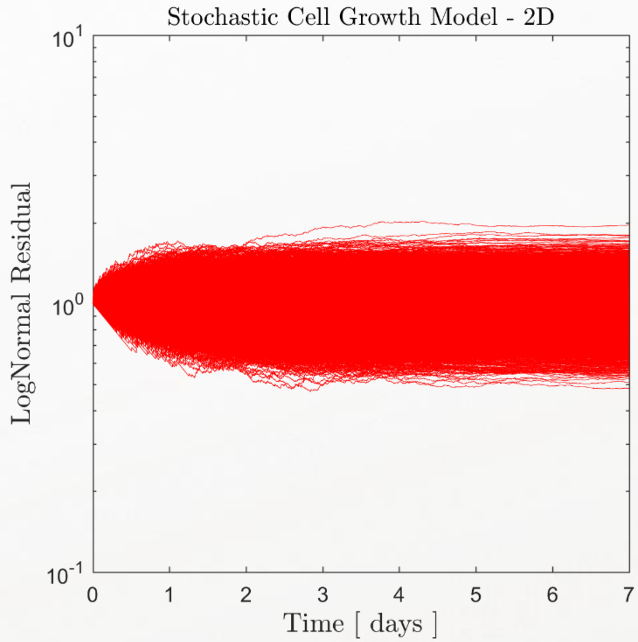
Central Limit Theorem

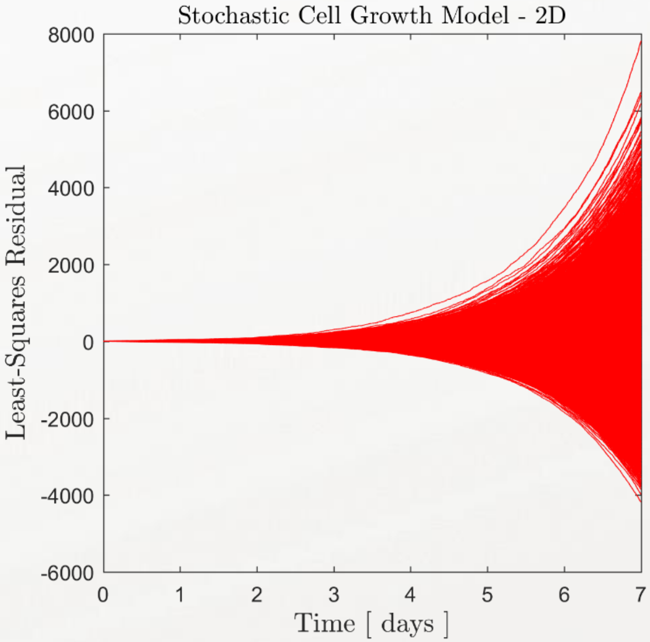
Sum of n independent random variables tends to normal distribution as $n \to \infty$.

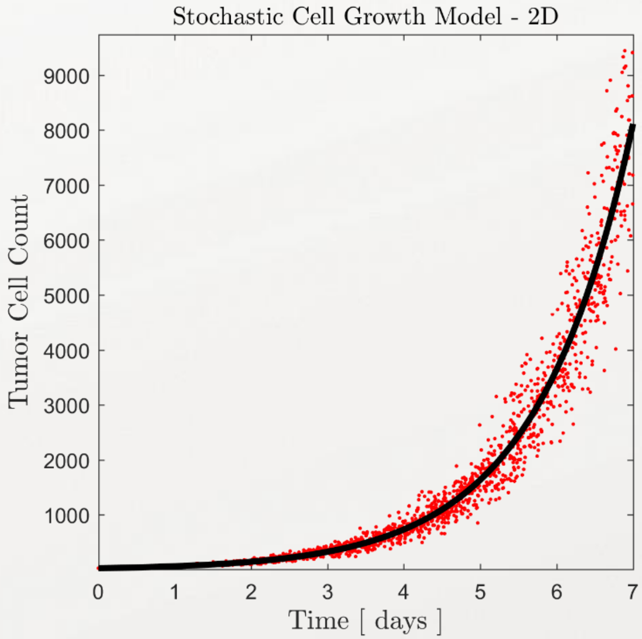
$$\Rightarrow \sum_{i=1}^{j} \ln \lambda_i \quad \sim \quad \mathcal{N}\big(\overline{\ln \lambda}, \sigma\big) \ ,$$

$$\Rightarrow \prod_{i=1}^{j} \lambda_i \quad \sim \quad \mathcal{LN}\big(\overline{\ln \lambda}, \sigma\big) \ .$$

Tumor growth modeling requires inadequacy models other than Least-Squares







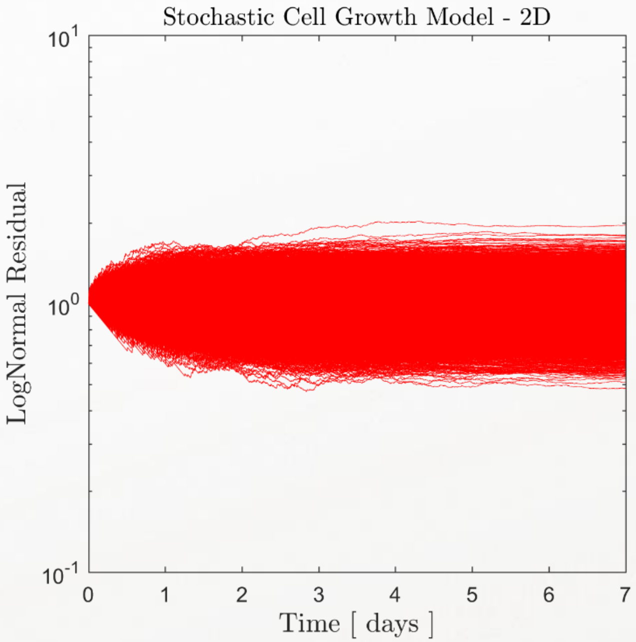
Central Limit Theorem

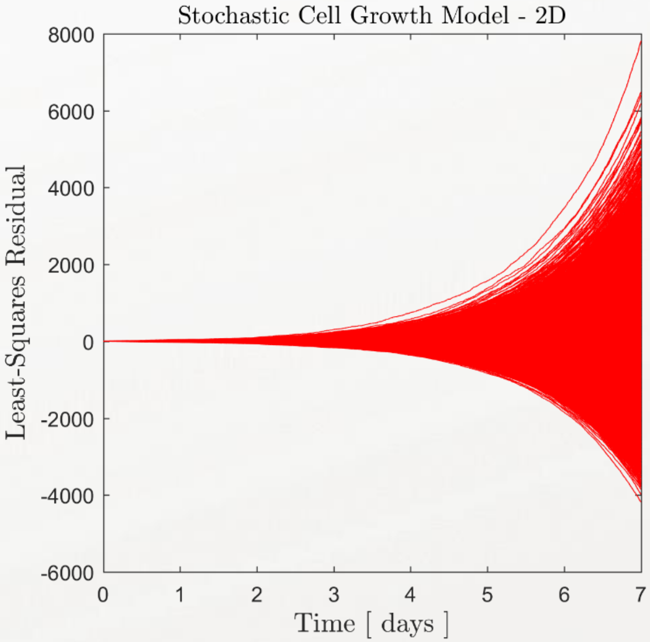
Sum of n independent random variables tends to normal distribution as $n \to \infty$.

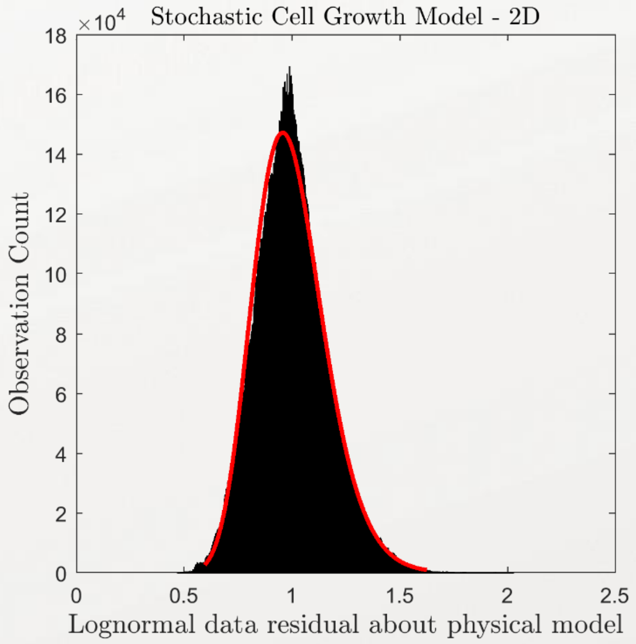
$$\Rightarrow \sum_{i=1}^{j} \ln \lambda_i \quad \sim \quad \mathcal{N}\left(\overline{\ln \lambda}, \sigma\right) \ ,$$

$$\Rightarrow \prod_{i=1}^{j} \lambda_i \quad \sim \quad \mathcal{LN}\left(\overline{\ln \lambda}, \sigma\right) \ .$$

Tumor growth modeling requires inadequacy models other than Least-Squares







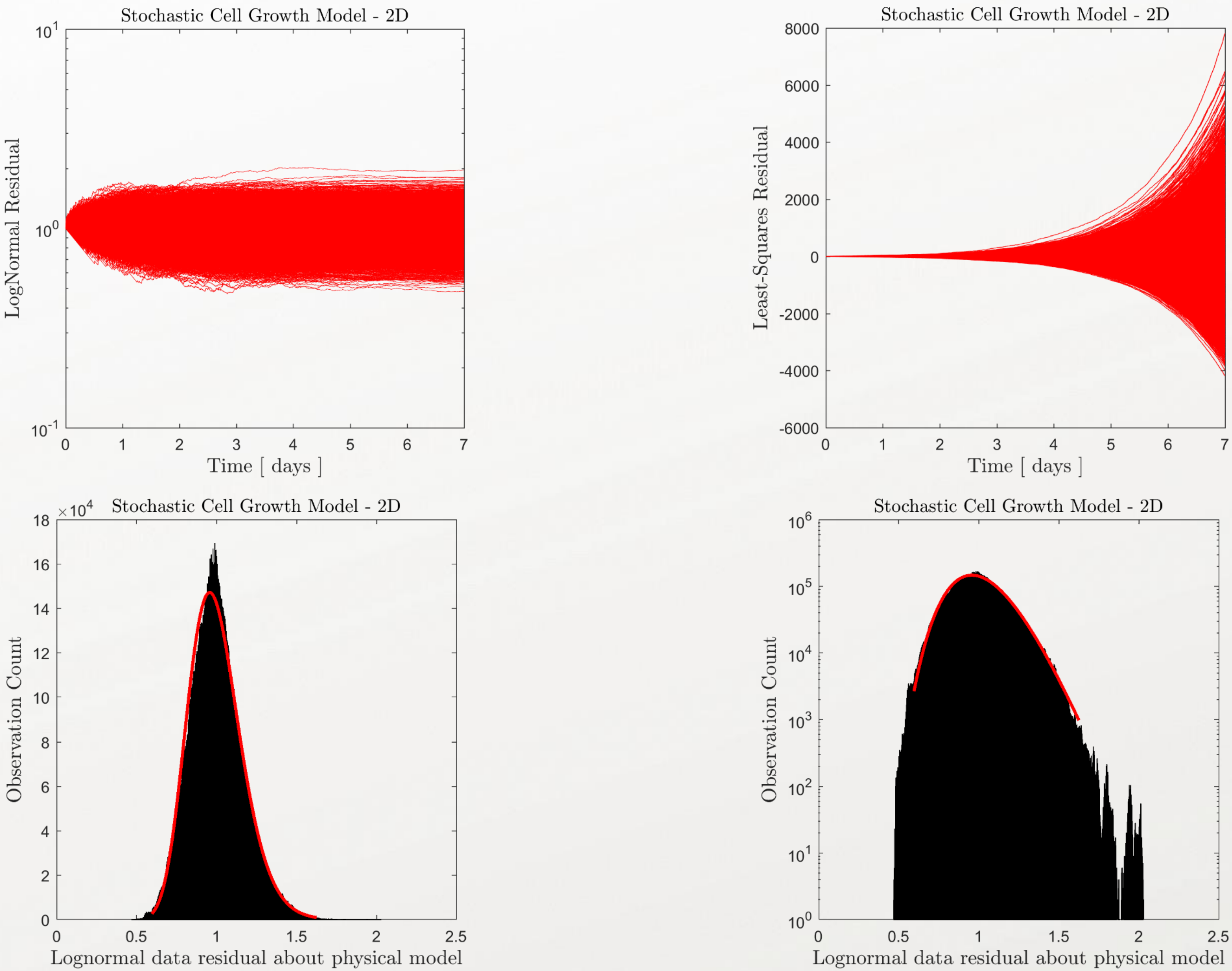
Central Limit Theorem

Sum of n independent random variables tends to normal distribution as $n \to \infty$.

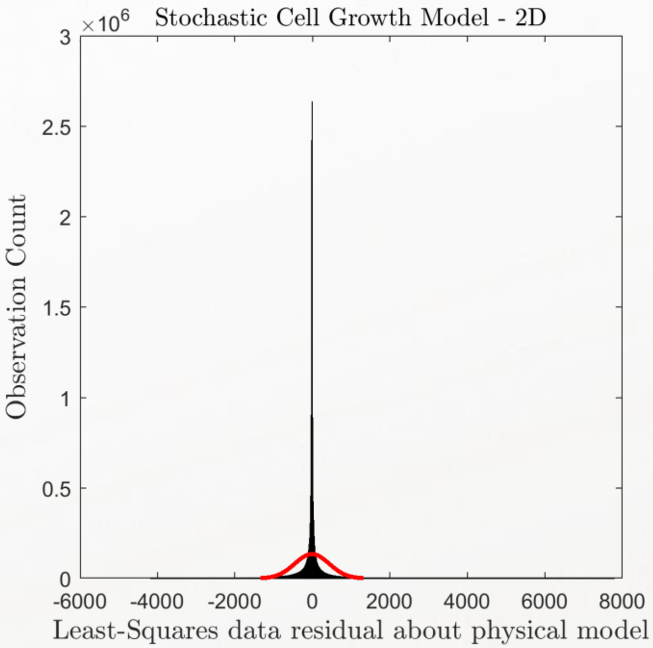
$$\Rightarrow \sum_{i=1}^{j} \ln \lambda_i \quad \sim \quad \mathcal{N}\left(\overline{\ln \lambda}, \sigma\right) \ ,$$

$$\Rightarrow \prod_{i=1}^{j} \lambda_i \quad \sim \quad \mathcal{LN}\left(\overline{\ln \lambda}, \sigma\right) \ .$$

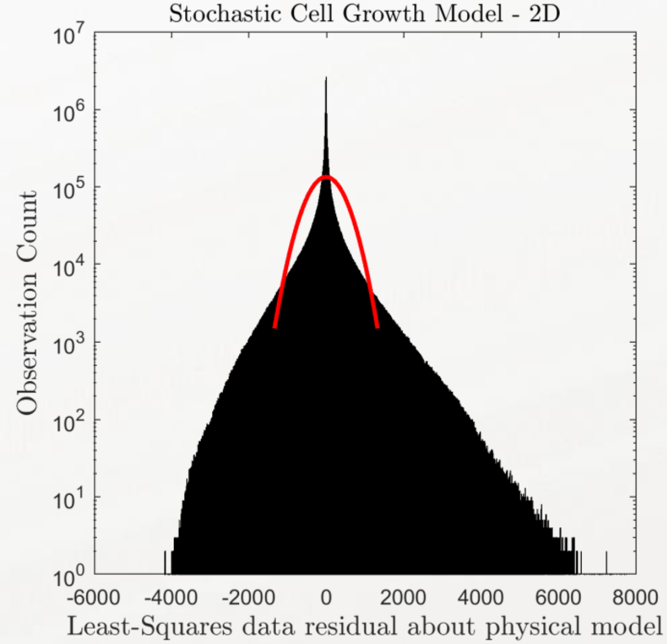
Tumor growth modeling requires inadequacy models other than Least-Squares

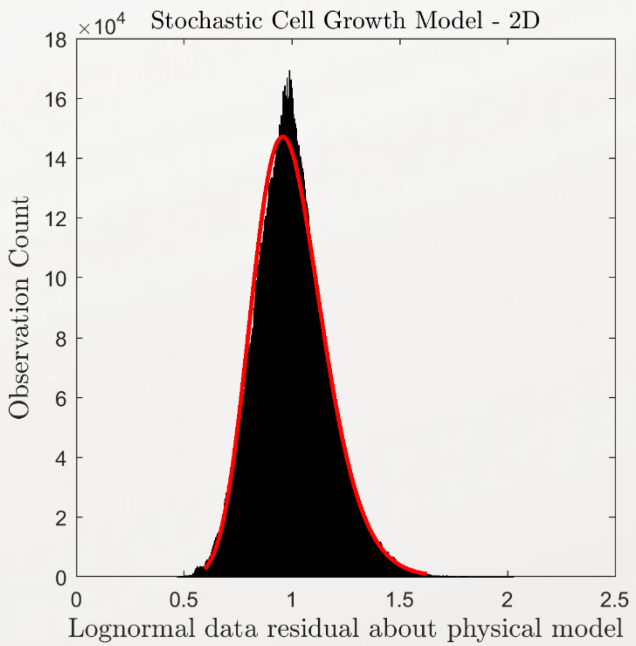


Tumor growth modeling requires inadequacy models other than Least-Squares

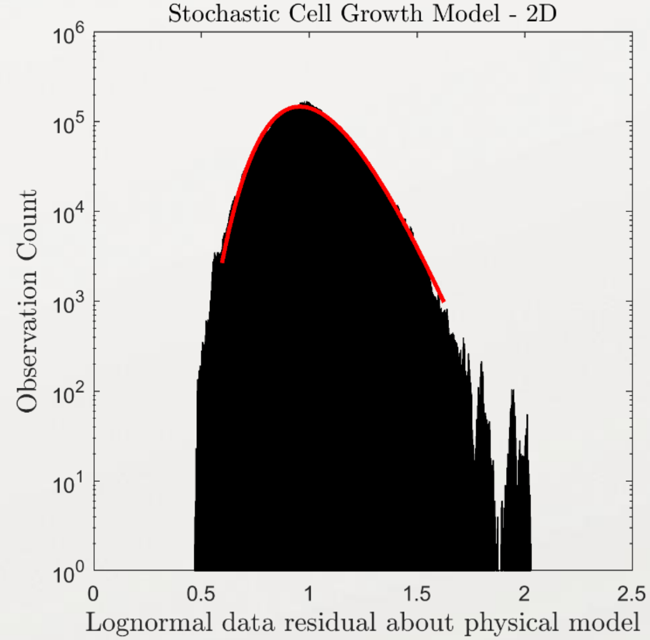


Normal likelihood (Least-Squares)





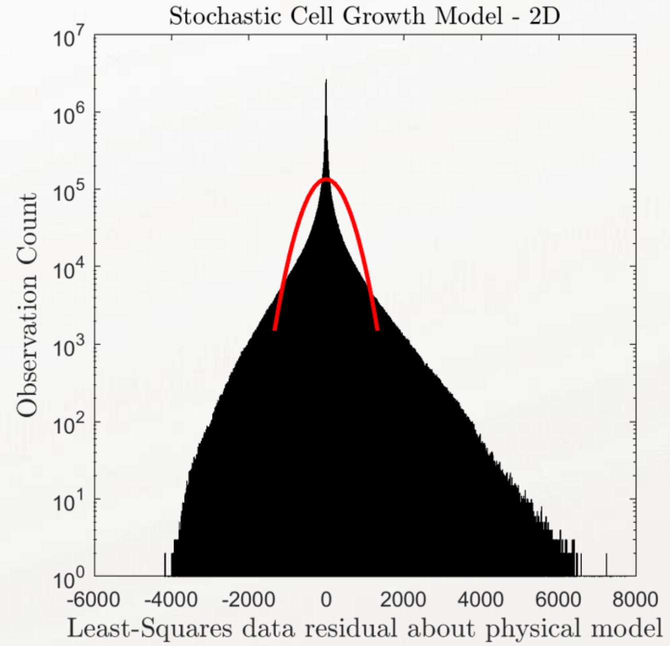
Lognormal Likelihood

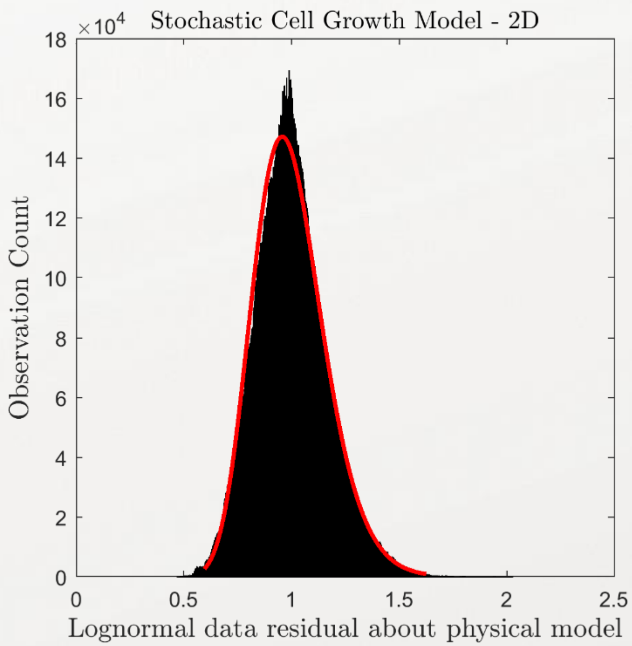


Why is Least-Squares method still so popular?

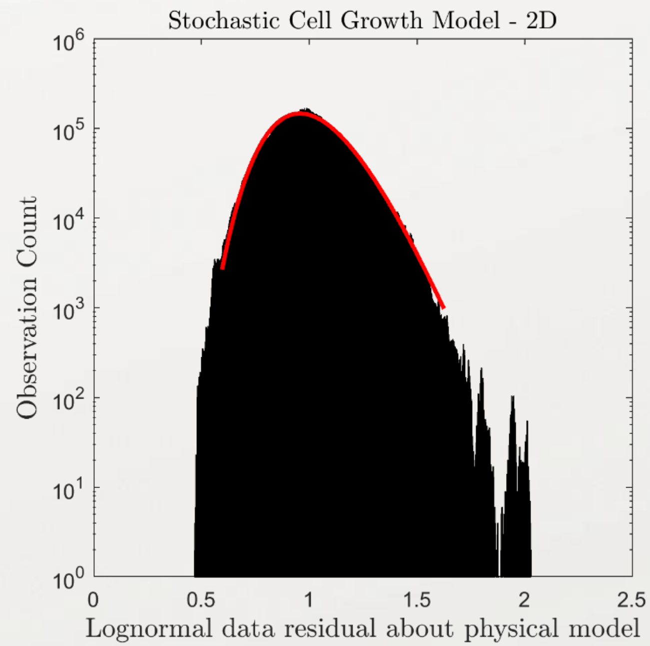
1. Many natural phenomena are normally distributed.
2. Lack of knowledge about its potential flaws (scarce data).
3. Confusion with noise.

Normal likelihood (Least-Squares)





Lognormal Likelihood



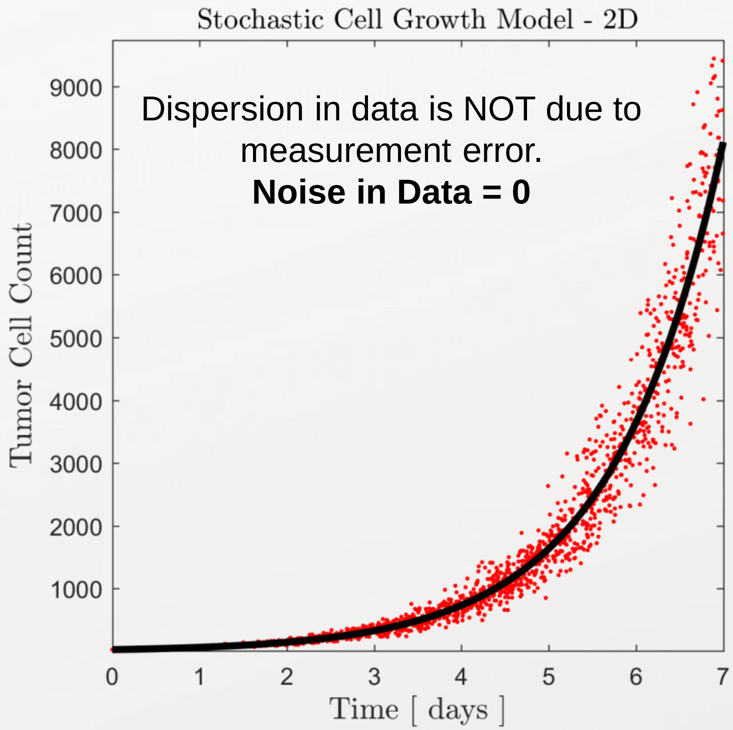
Why is Least-Squares method still so popular?

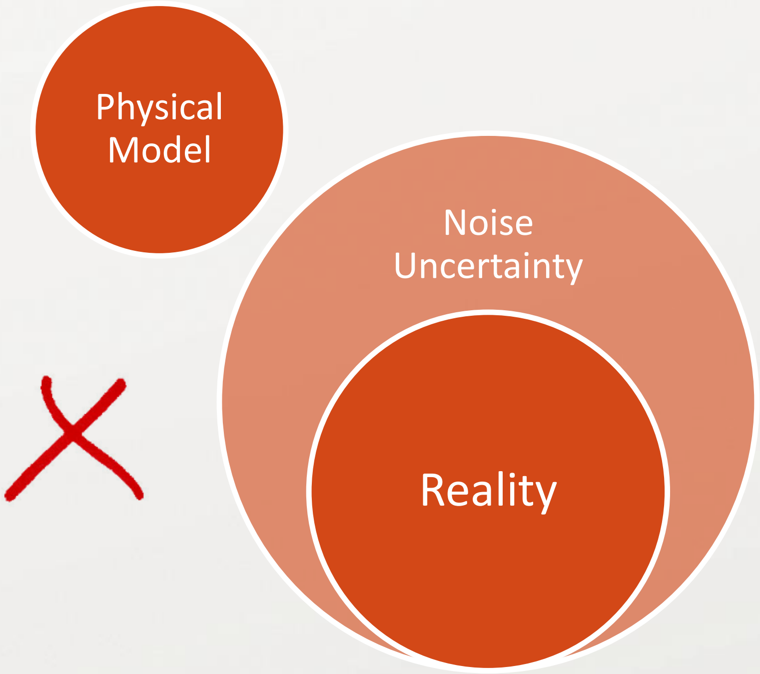
1. Many natural phenomena are normally distributed.
2. Lack of knowledge about its potential flaws (scarce data).
3. Confusion with noise.

Consensus likelihood function

$$\pi\left(\text{Dataset}\middle|\text{Model}\right) \propto \prod_{i=1}^{n_{\text{do}}} \exp\left(- \left[\frac{\text{Data[i]} - \text{Model}(\ \text{Data[i]}\ )}{\text{Noise in Data[i]}}\right]^2 \right)$$

Strong underlying assumption: The physical **model** is **perfectly** and **deterministically correct**.
No need for **model inadequacy**.
All **uncertainty is** due to **measurement error**.





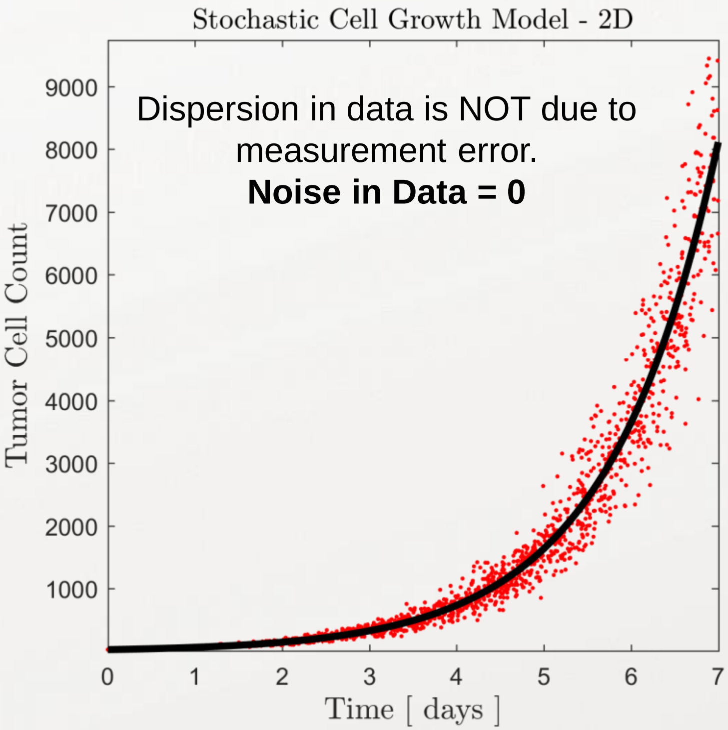
Why is Least-Squares method still so popular?

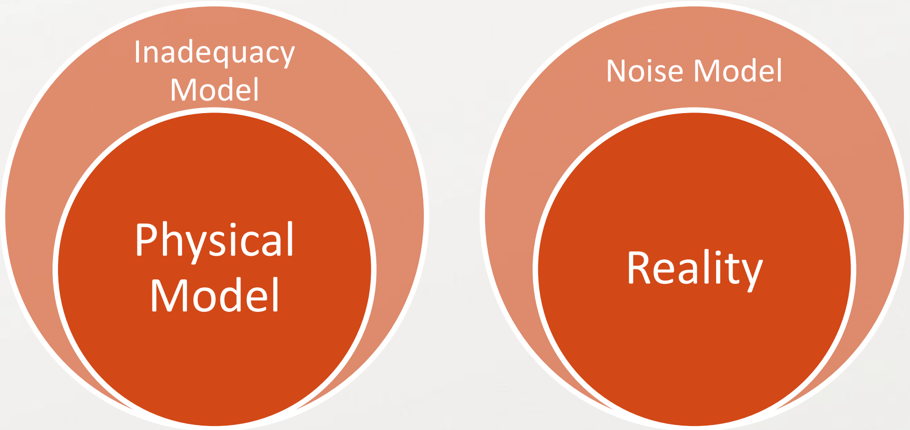
1. Many natural phenomena are normally distributed.
2. Lack of knowledge about its potential flaws (scarce data).
3. Confusion with noise.

Consensus likelihood function

$$\pi\left(\text{Dataset}\,\middle|\,\text{Model}\right) \propto \prod_{\text{i}=1}^{n_{\text{do}}} \exp\left(- \left[\frac{\text{Data[i]} - \text{Model}\left(\ \text{Data[i]}\ \right)}{\text{Noise in Data[i]}} \right]^2 \right)$$

Strong underlying assumption: The physical **model** is **perfectly** and **deterministically correct**.
No need for **model inadequacy**.
All **uncertainty is** due to **measurement error**.

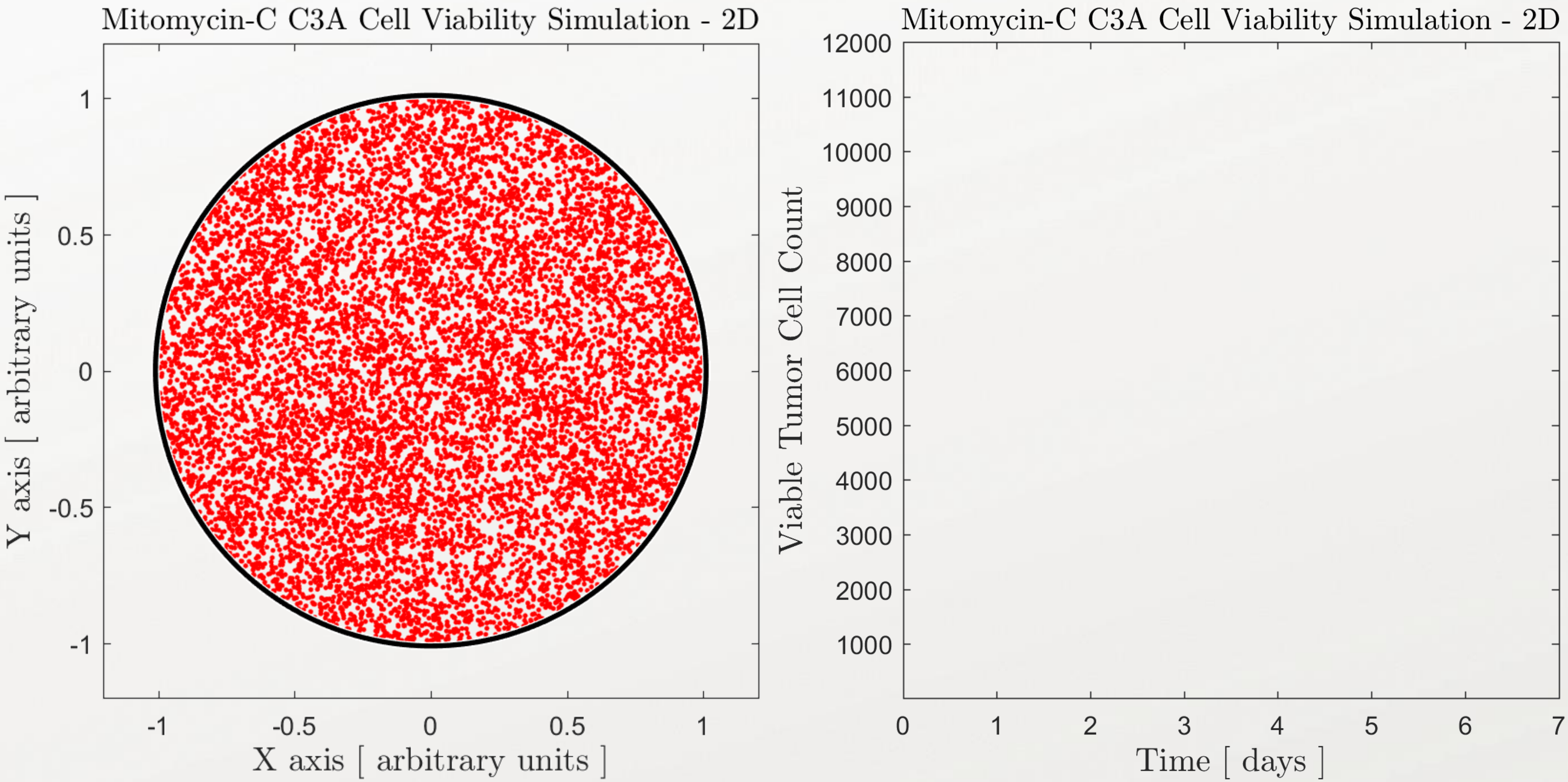




Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

The perfect deterministic physical model:

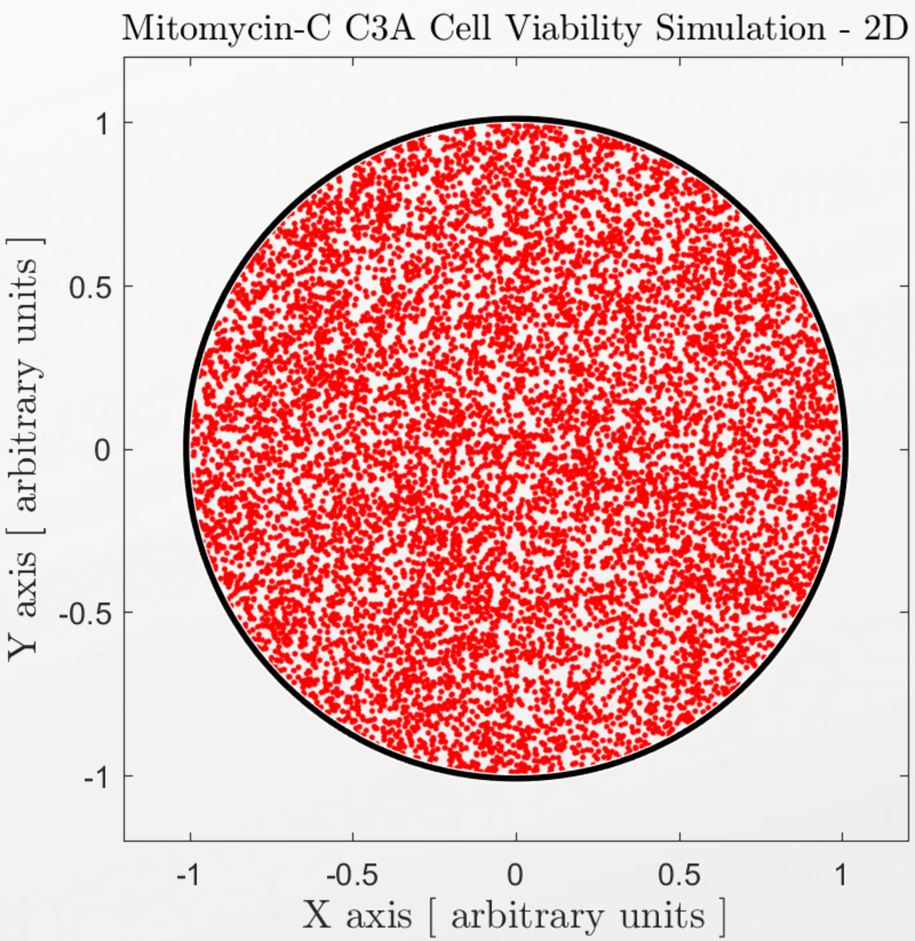
$$\mathrm{d}\,N = -\lambda N \,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{-\lambda t}$$

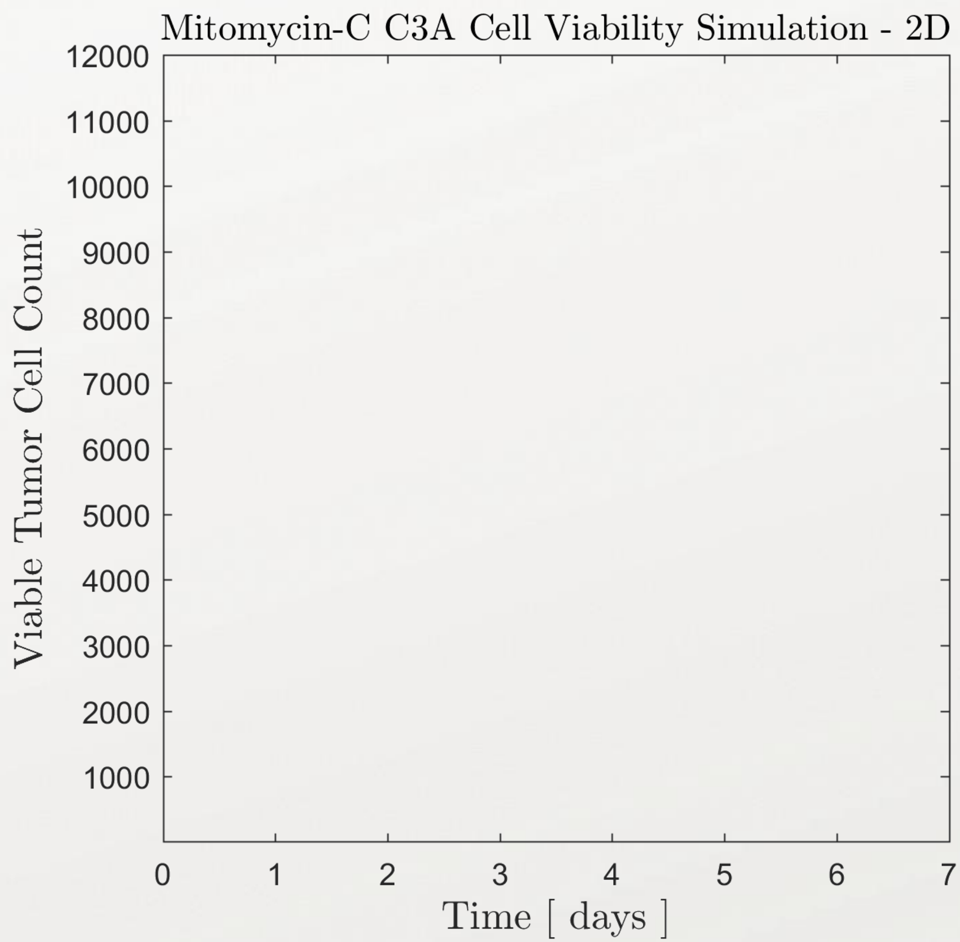


Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

The perfect deterministic physical model:

$$\mathrm{d}\,N = -\lambda N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{-\lambda t}$$

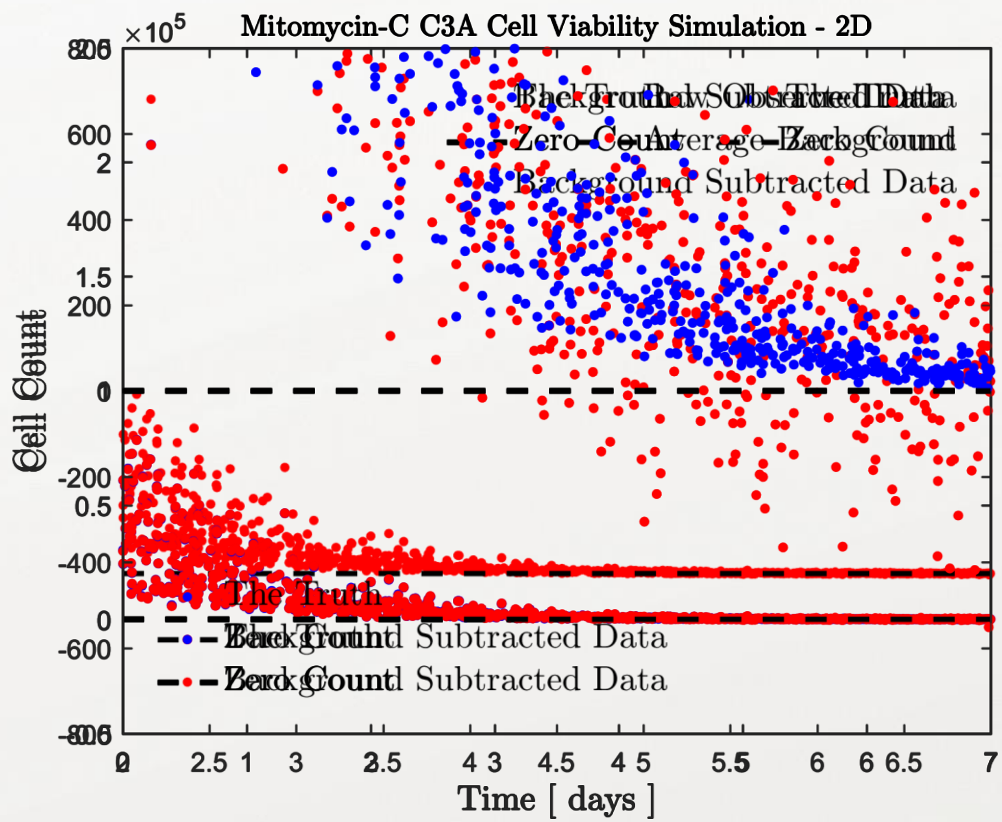




Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

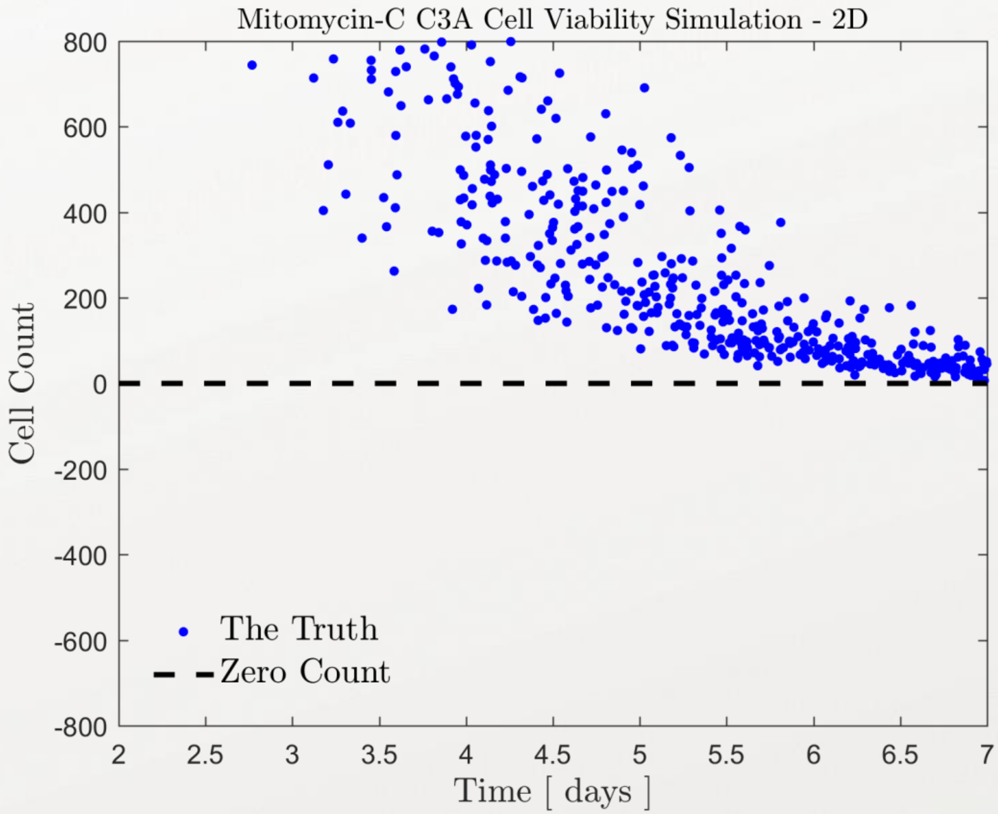
The perfect deterministic physical model:

$$\mathrm{d}\,N = -\lambda N\,\mathrm{d}\,t \quad \Rightarrow \quad N(t) = N_0\,\mathrm{e}^{-\lambda t}$$



Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

Physical Model + Inadequacy Model + Noise Model $\Rightarrow$ Infer the unknown Truth



Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Reality (truth)

$$\mathcal{R} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_{\mathrm{ndo}}\}$$

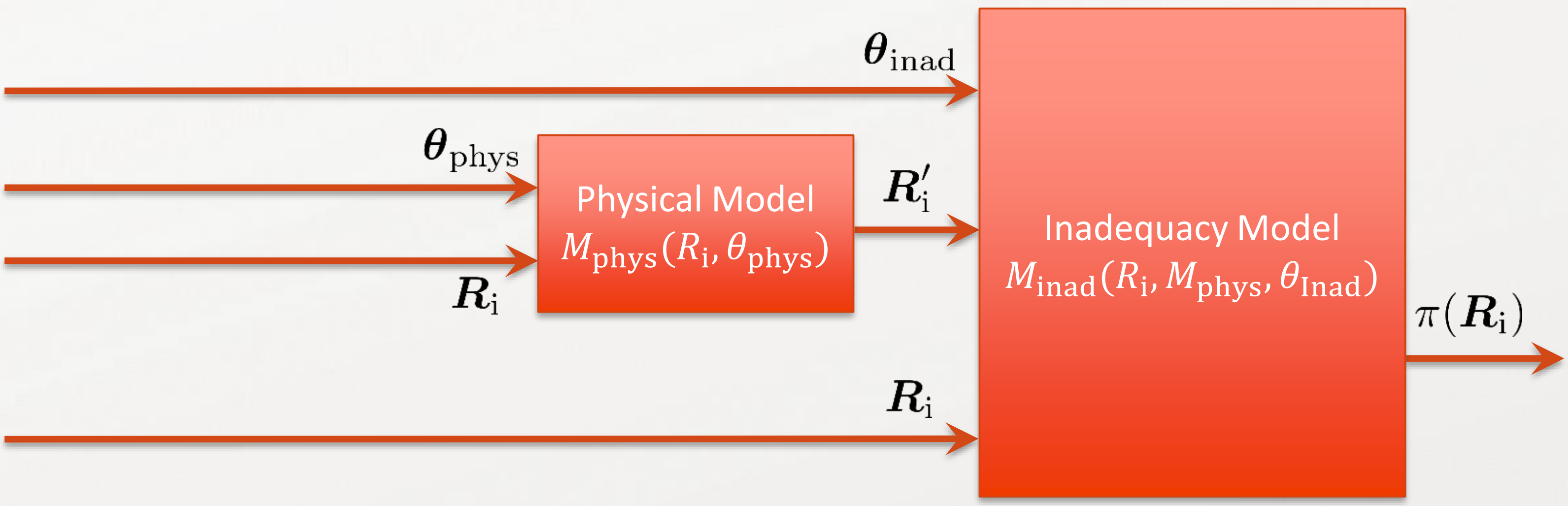
$$\boldsymbol{D}_{\mathrm{i}} \sim \boldsymbol{M}_{\mathrm{nois,i}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}})$$

Data (subject to noise)

$$\mathcal{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_{\mathrm{ndo}}\}$$

$\mathcal{R}$	$\boldsymbol{X}$	$\boldsymbol{Y}$	$\boldsymbol{Z}$	$\boldsymbol{t}$	$\boldsymbol{N}$
$\boldsymbol{R_1}$	x_1	y_1	z_1	t_1	N_1
$\boldsymbol{R_2}$	x_2	y_2	z_2	t_2	N_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{R_{n_{do}}}$	$x_{n_{do}}$	$y_{n_{do}}$	$z_{n_{do}}$	$t_{n_{do}}$	$N_{n_{do}}$

$\mathcal{D}$	$\boldsymbol{X}$	$\boldsymbol{Y}$	$\boldsymbol{Z}$	$\boldsymbol{t}$	$\boldsymbol{N}$
$\boldsymbol{D_1}$	x'_1	y'_1	z'_1	t'_1	N'_1
$\boldsymbol{D_2}$	x'_2	y'_2	z'_2	t'_2	N'_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{D_{n_{do}}}$	$x'_{n_{do}}$	$y'_{n_{do}}$	$z'_{n_{do}}$	$t'_{n_{do}}$	$N'_{n_{do}}$



Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Reality (truth)

$$\mathcal{R} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_{\mathrm{n_{do}}}\}$$

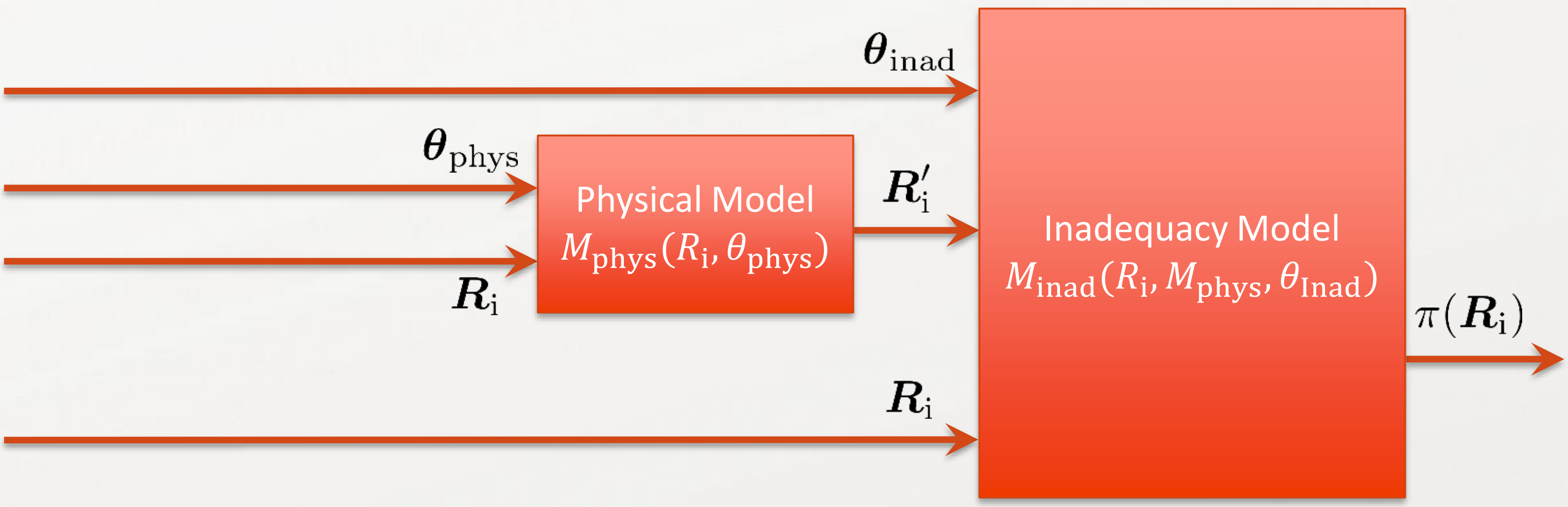
$$\boldsymbol{D}_{\mathrm{i}} \sim \boldsymbol{M}_{\mathrm{nois,i}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}})$$

Data (subject to noise)

$$\mathcal{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_{\mathrm{n_{do}}}\}$$

$$\mathrm{PDF}(\boldsymbol{D}_{\mathrm{i}}) \stackrel{\mathrm{def}}{=} \pi\left(\boldsymbol{D}_{\mathrm{i}} \middle| \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}\right)$$

$$\pi\left(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} \;\middle|\; \boldsymbol{D}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}, \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}}\right) = \frac{\pi\left(\boldsymbol{D}_{\mathrm{i}} \middle| \boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}\right) \pi\left(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} \middle| \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}}\right)}{\pi\left(\boldsymbol{D}_{\mathrm{i}} \middle| \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}\right)}$$



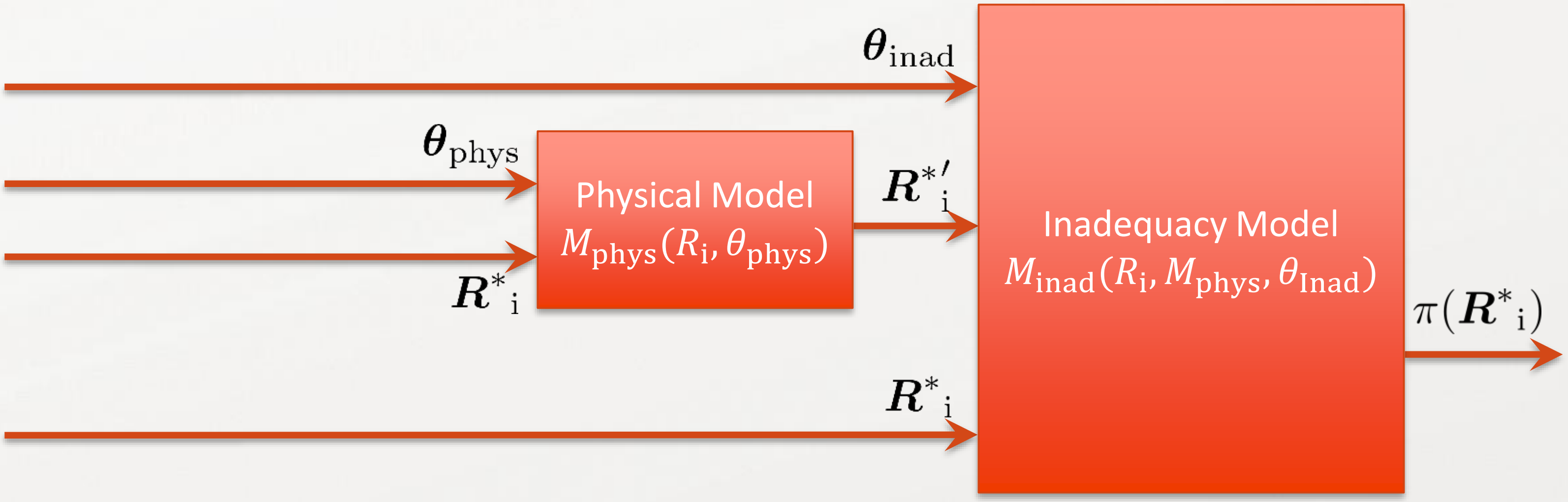
Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Reality (truth) **Data (subject to noise)**

$$\mathcal{R} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_{\mathrm{n_{do}}}\} \xrightarrow{\boldsymbol{D}_{\mathrm{i}} \sim \boldsymbol{M}_{\mathrm{nois,i}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}})} \mathcal{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_{\mathrm{n_{do}}}\}$$

$$\mathrm{PDF}(\boldsymbol{D}_{\mathrm{i}}) \stackrel{\mathrm{def}}{=} \pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})$$

$$\pi(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} \mid \boldsymbol{D}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}, \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}}) = \frac{\pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})\ \pi(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} | \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}})}{\pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})}$$



Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Reality (truth)

$$\mathcal{R} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, ..., \boldsymbol{R}_{\mathrm{n_{do}}}\}$$

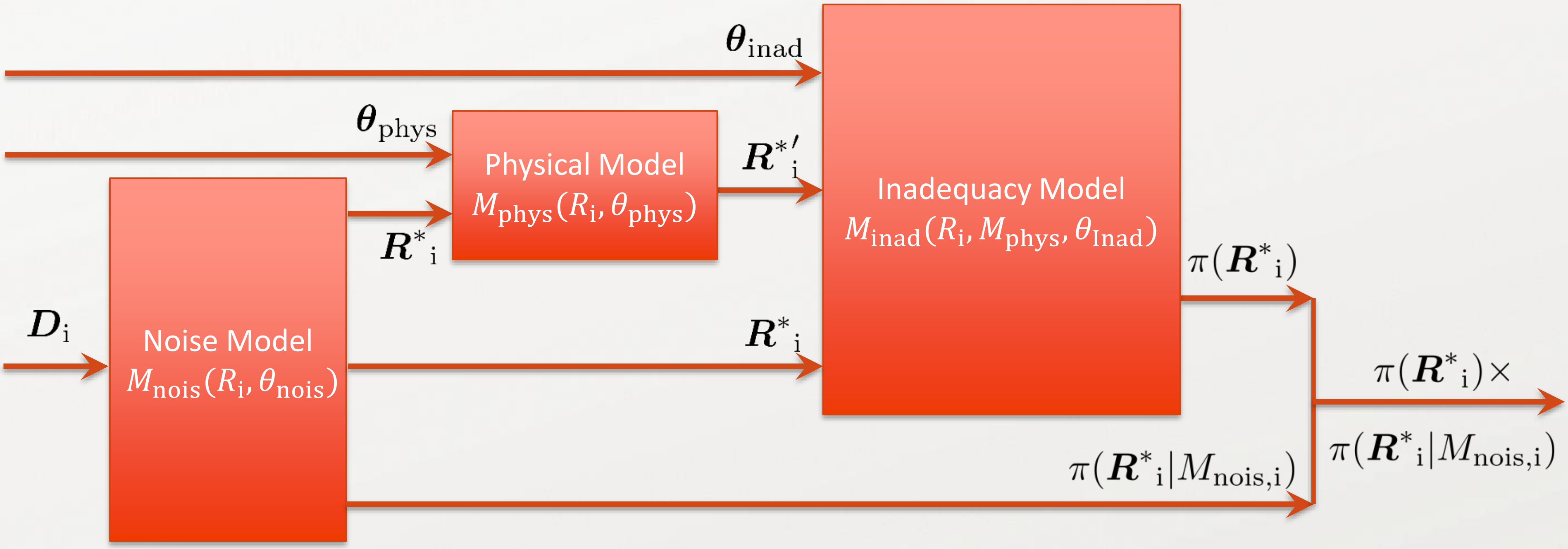
$$\boldsymbol{D}_{\mathrm{i}} \sim \boldsymbol{M}_{\mathrm{nois,i}}(\boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}})$$

Data (subject to noise)

$$\mathcal{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, ..., \boldsymbol{D}_{\mathrm{n_{do}}}\}$$

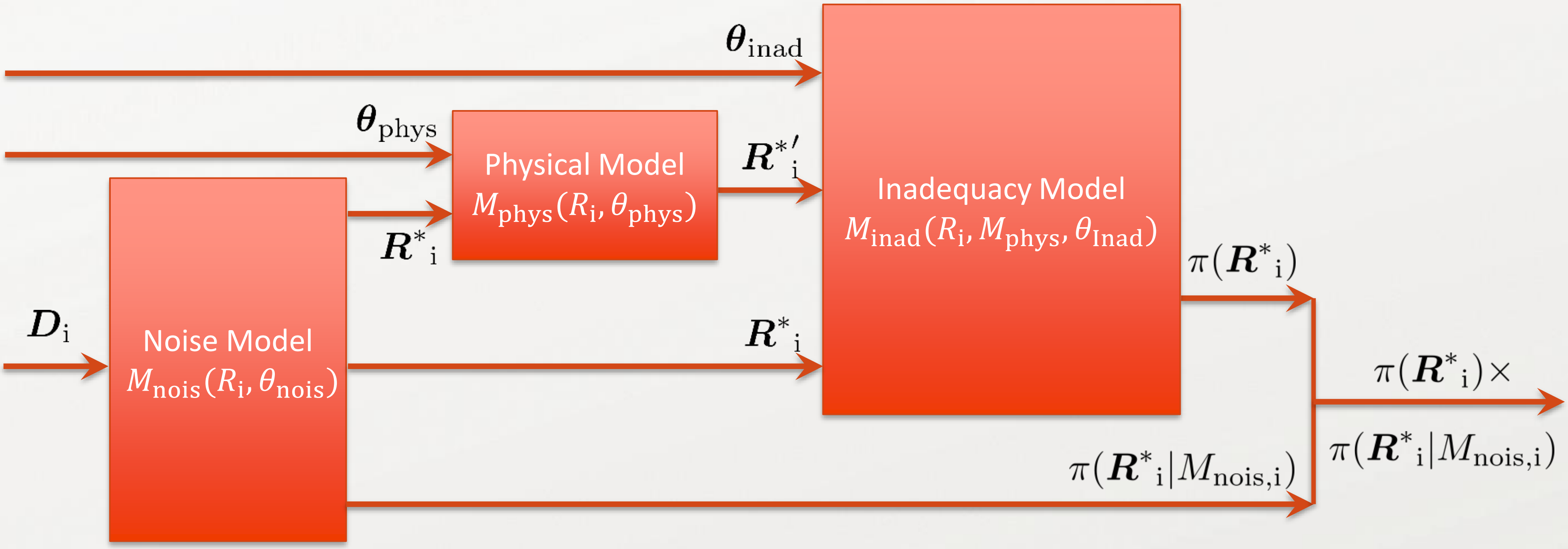
$$\mathrm{PDF}(\boldsymbol{D}_{\mathrm{i}}) \stackrel{\mathrm{def}}{=} \pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})$$

$$\pi(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} \mid \boldsymbol{D}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}}, \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}}) = \frac{\pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}}, \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})\ \pi(\boldsymbol{R}^*{}_{\mathrm{i}} = \boldsymbol{R}_{\mathrm{i}} | \mathcal{I}_{\boldsymbol{R}_{\mathrm{i}}})}{\pi(\boldsymbol{D}_{\mathrm{i}} | \boldsymbol{\theta}_{\mathrm{nois,i}}, \boldsymbol{M}_{\mathrm{nois,i}})}$$

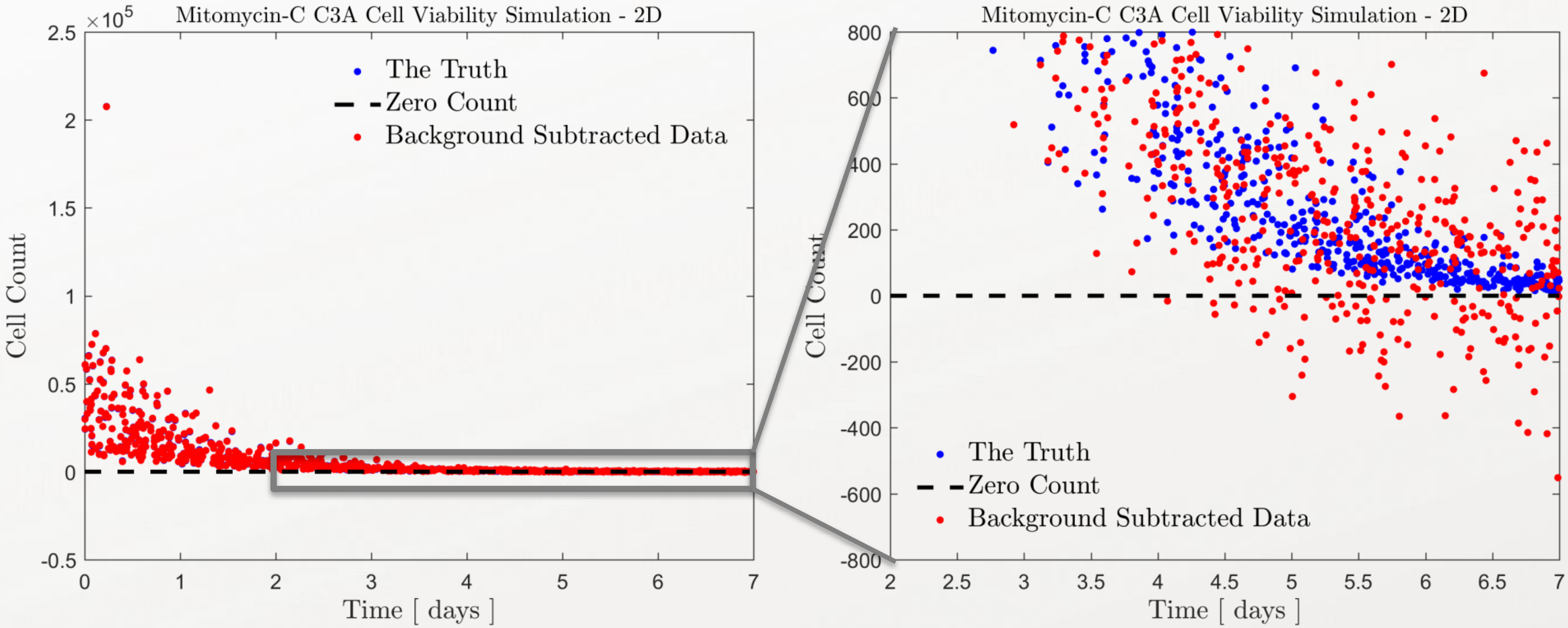


Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

$$\begin{aligned}\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) &= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}} \;;\; \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\ &= \frac{\int_{\boldsymbol{\Omega}_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\; \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\, \mathrm{d}\mathcal{R}^*\; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)},\end{aligned}$$



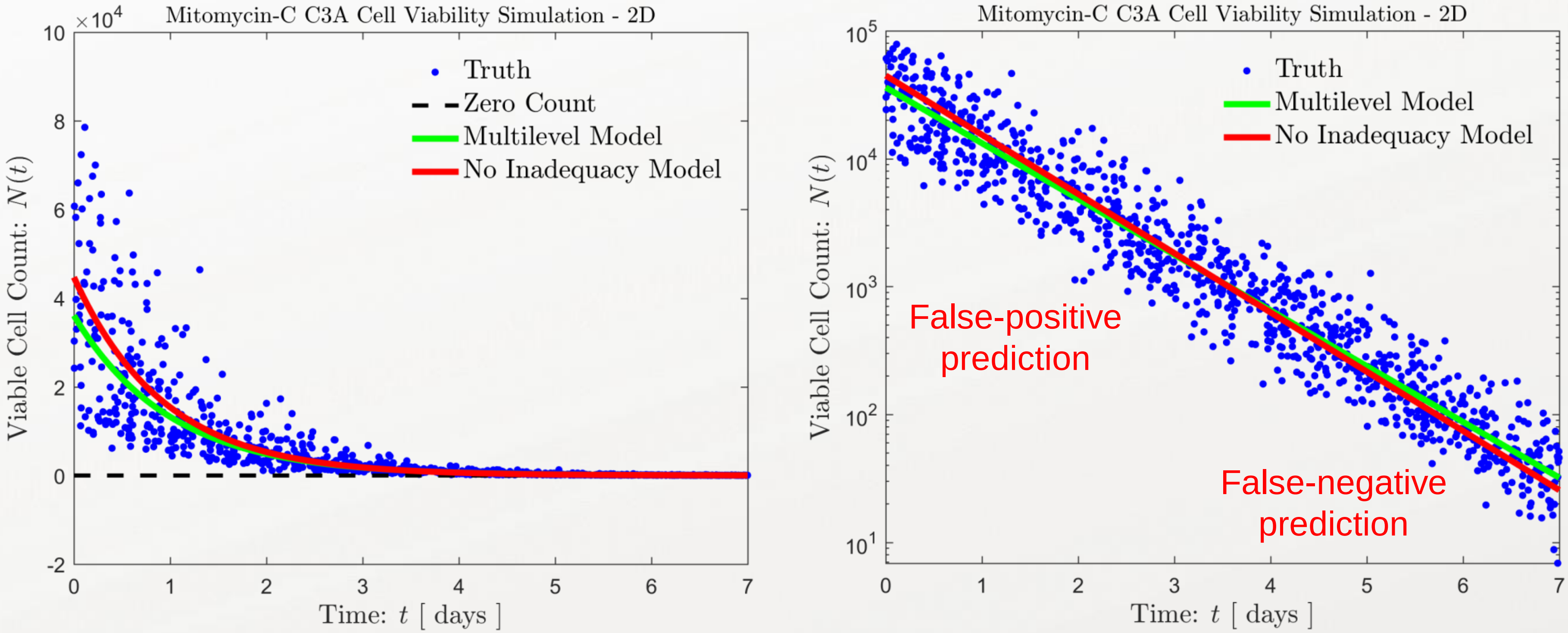
Multilevel Bayesian modeling of data in the presence of model inadequacy and noise



Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Consensus likelihood function

$$\pi\left(\text{Dataset}\,\middle|\,\text{Model}\right) \propto \prod_{\text{i}=1}^{n_{\text{do}}} \exp\left(- \left[\frac{\text{Data[i]} - \text{Model}\left(\ \text{Data[i]}\ \right)}{\text{Noise in Data[i]}} \right]^2 \right)$$



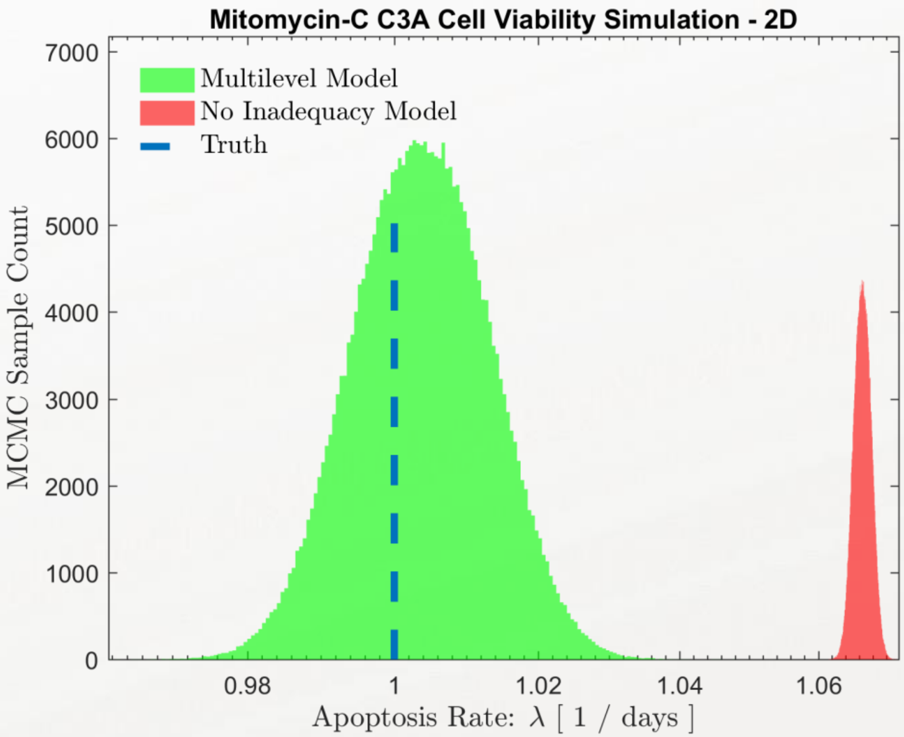
Multilevel Model

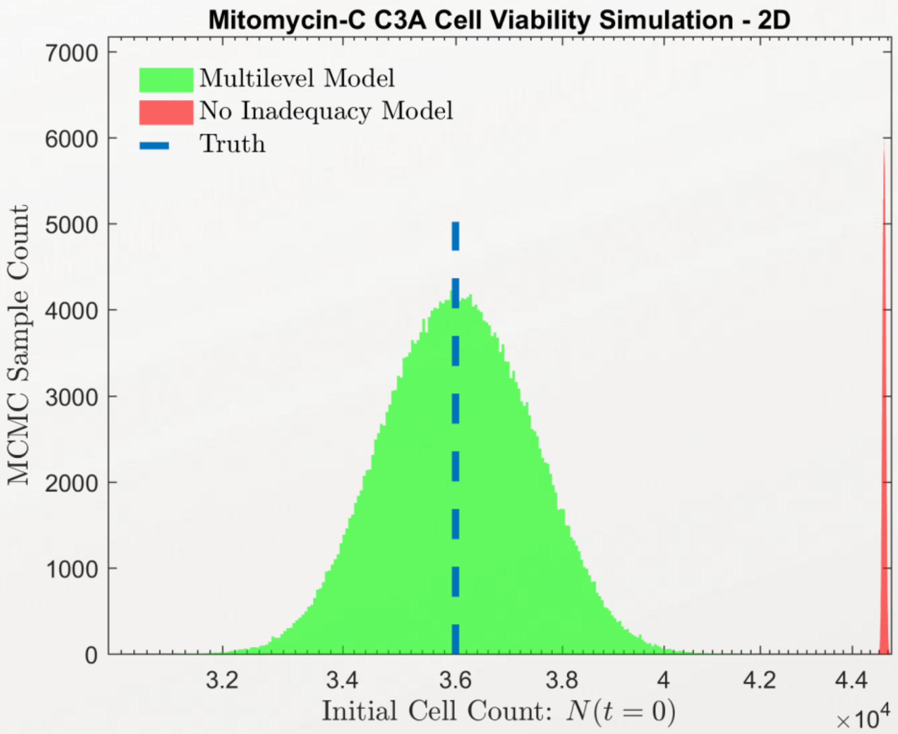
$$\begin{aligned}
\pi\left(\boldsymbol{\theta}_{\text{pi}} \,\middle|\, \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)
&= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\text{pi}}\ ;\ \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \pi\left(\boldsymbol{\theta}_{\text{pi}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)} \\
&= \frac{\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\text{pi}}, \boldsymbol{M}_{\text{pi}}\right)\ \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \mathrm{d}\mathcal{R}^*\ \pi\left(\boldsymbol{\theta}_{\text{pi}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}
\end{aligned}$$

Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Consensus likelihood function

$$\pi(\mathrm{Dataset}|\mathrm{Model}) \propto \prod_{\mathrm{i}=1}^{n_{\mathrm{do}}} \exp\left(- \left[\frac{\mathrm{Data[i]} - \mathrm{Model(\ Data[i]\ )}}{\mathrm{Noise\ in\ Data[i]}} \right]^2 \right)$$



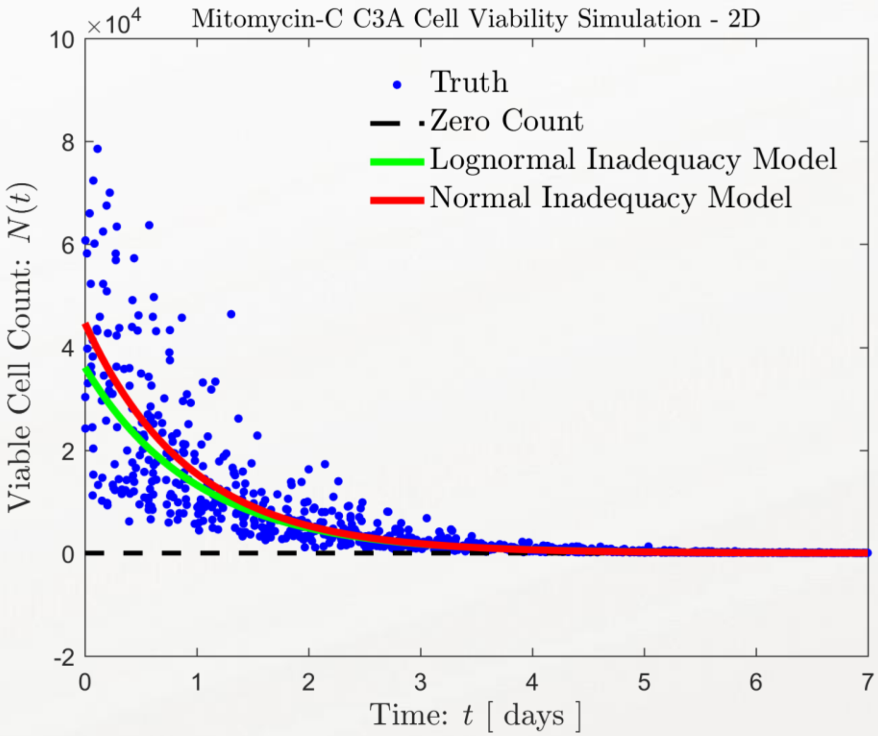


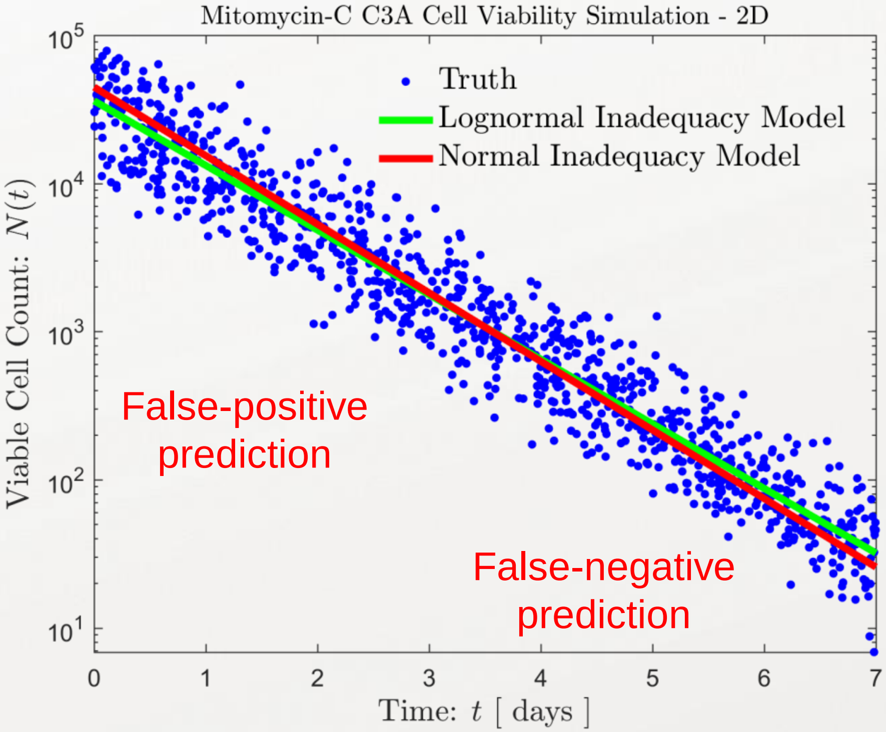
Multilevel Model

$$\begin{aligned}\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) &= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}}\ ;\ \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} | \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} | \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\ &= \frac{\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^* | \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\ \pi\left(\mathcal{R}^* | \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \mathrm{d}\mathcal{R}^*\ \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} | \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} | \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}\end{aligned}$$

Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

How about the effects of wrong inadequacy model?





Multilevel Model

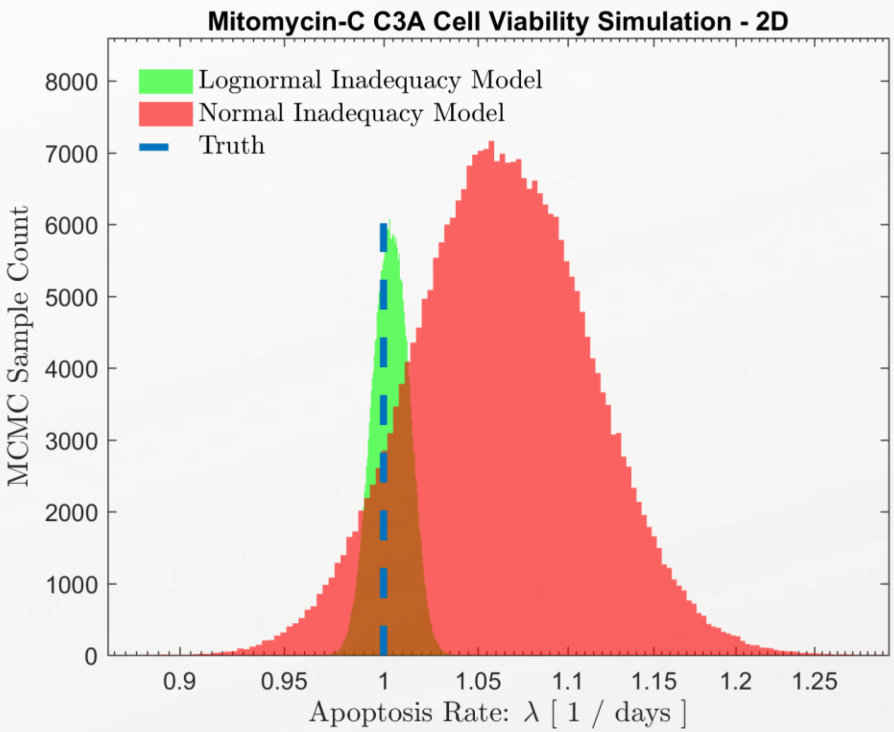
$$\pi\left(\boldsymbol{\theta}_{\text{pi}} \;\middle|\; \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)$$

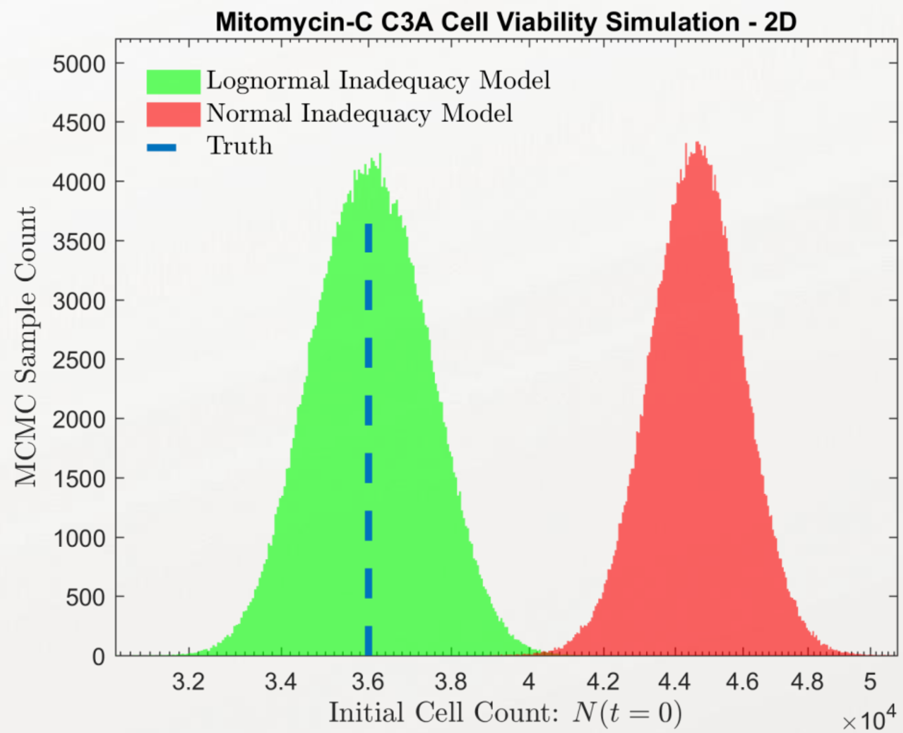
$$= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\text{pi}} \; ; \; \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \mathcal{I}_{\mathcal{R}}\right) \; \pi\left(\boldsymbol{\theta}_{\text{pi}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}$$

$$= \frac{\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\text{pi}}, \boldsymbol{M}_{\text{pi}}\right) \; \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}}, \mathcal{I}_{\mathcal{R}}\right) \; \mathrm{d}\mathcal{R}^* \; \pi\left(\boldsymbol{\theta}_{\text{pi}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\text{nois}}, \mathcal{M}_{\text{nois}} \middle| \boldsymbol{M}_{\text{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\text{pi}}}\right)}$$

Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

How about the effects of wrong inadequacy model?



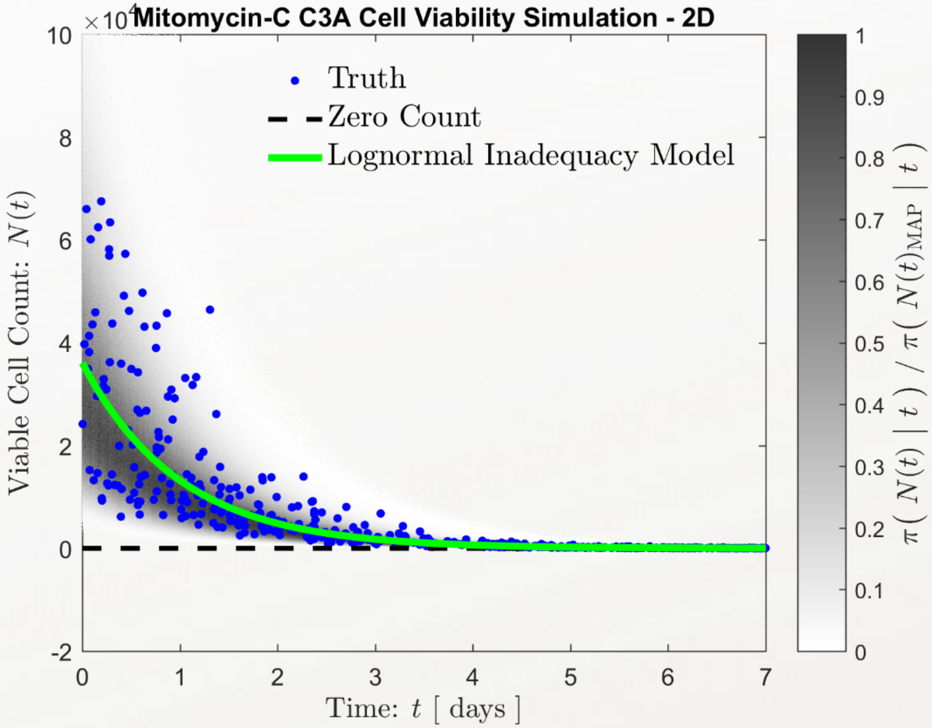


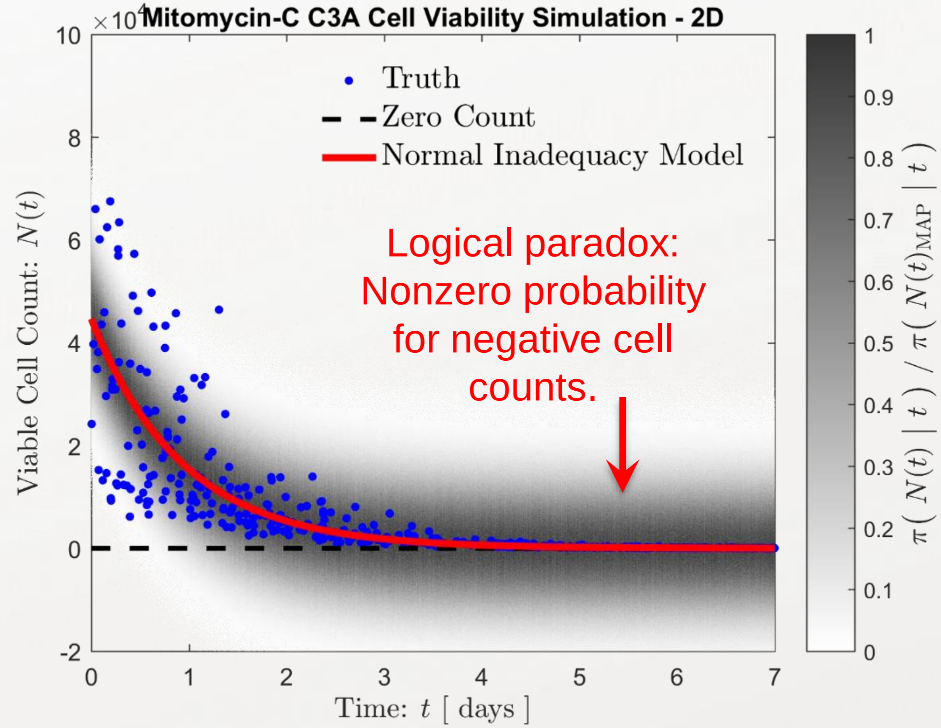
Multilevel Model

$$
\begin{aligned}
\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)
&= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}} \;;\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\
&= \frac{\displaystyle\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^{*} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \; \pi\left(\mathcal{R}^{*} \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \, \mathrm{d}\mathcal{R}^{*} \; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}
\end{aligned}
$$

Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

How about the effects of wrong inadequacy model?



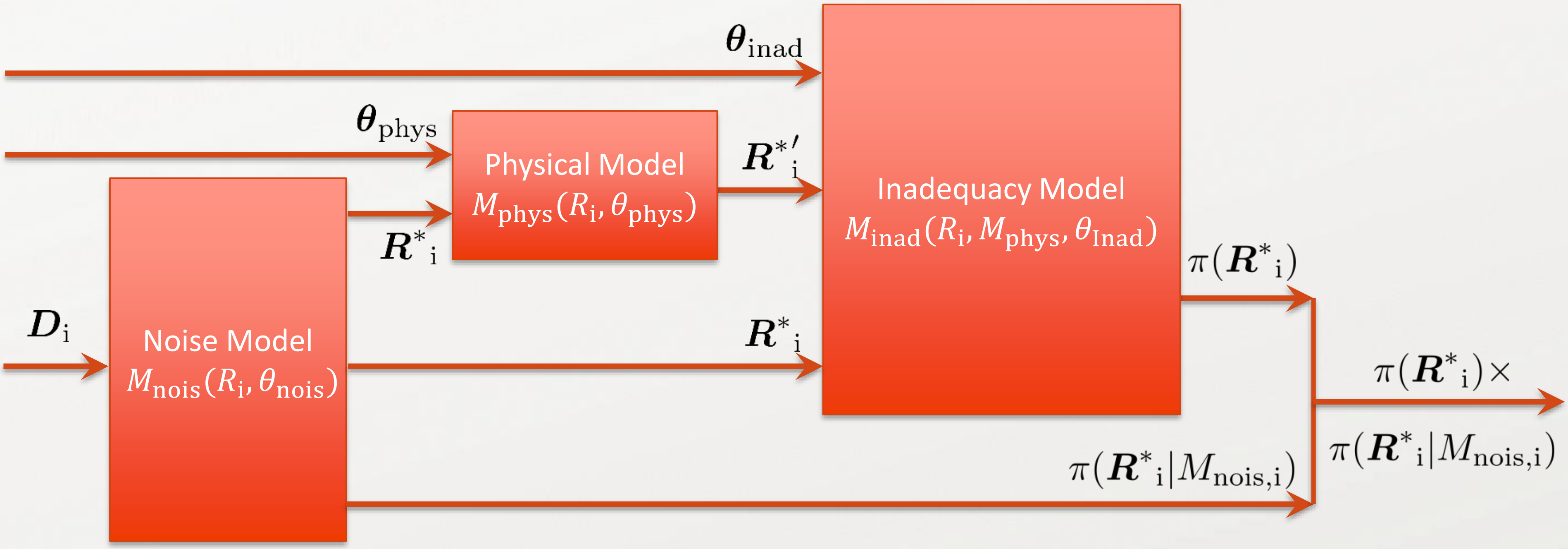


Multilevel Model

$$
\begin{aligned}
\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)
&= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}} \;;\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\
&= \frac{\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \mathrm{d}\mathcal{R}^* \; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}
\end{aligned}
$$

Summary

- There is only one type of uncertainty in scientific inference: **epistemic** or **lack of knowledge.**
- Epistemic uncertainty can manifest itself in **two fundamentally different forms**:
    - Model inadequacy
    - Experimental noise
- Confusion of model inadequacy with noise or a wrong choice of inadequacy model can lead to
    - **Logical paradoxes**: negative number of tumor cells, negative concentration, …
    - Increased likelihood of **false-negative** and **false-positive** conclusions.
- The correct inadequacy model for tumor modeling is lognormal, not least-squares.

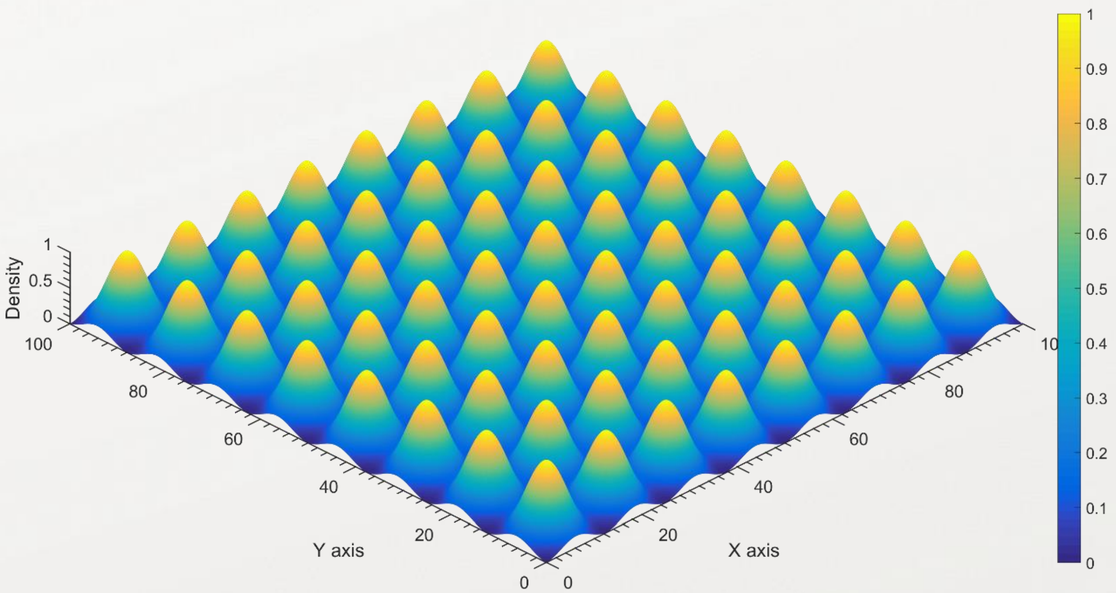


Lebesgue Monte Carlo Integration of Bayesian Evidence

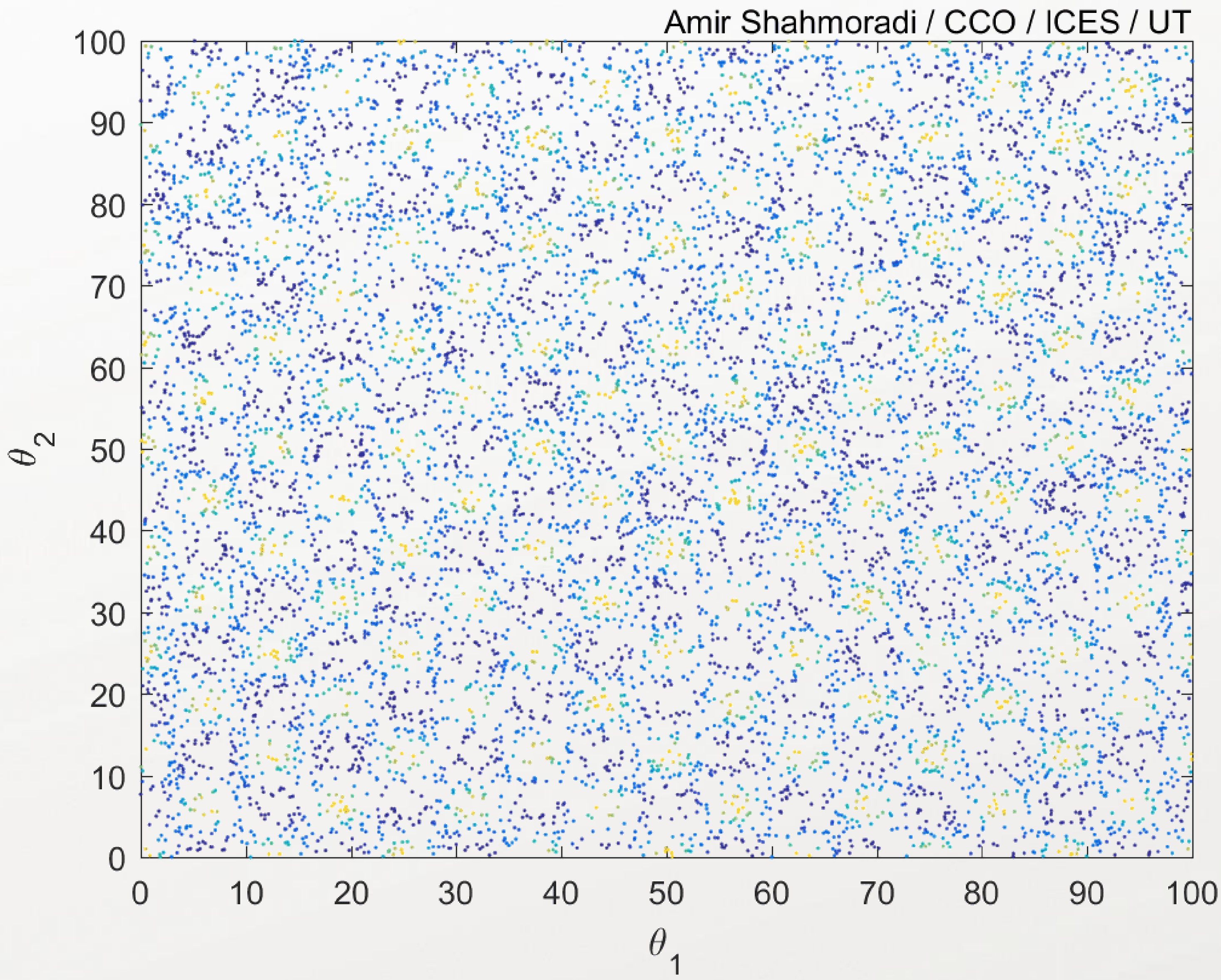
Physical model + inadequacy model + noise model:

$$
\begin{aligned}
&\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) \\
&\quad = \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}} \,;\, \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \, \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\
&\quad = \frac{\displaystyle\int_{\boldsymbol{\Omega}_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \, \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right) \, \mathrm{d}\mathcal{R}^* \; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \;,
\end{aligned}
$$

Amir Shahmoradi / CCO / ICES / UT



Lebesgue Monte Carlo Integration of Bayesian Evidence



Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Consensus likelihood function?

$$\ln\left(\pi(\boldsymbol{y}_c|\boldsymbol{\theta}_i,\mathcal{M}_i)\right) = \frac{N}{2}\ln(2\pi) + \sum_{j=1}^{N}\left[-\ln(\sigma_j) - \frac{1}{2}\left(\frac{y_{cj}-d_j^c(\boldsymbol{\theta}_i)}{\sigma_j}\right)^2\right]$$

Physical + noise models:

$$\pi\left(\boldsymbol{\theta}_{\mathrm{phys}} \mid \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right) = \frac{\left[\int_{\boldsymbol{\mathcal{R}}^{*}{}_{\theta_{\mathrm{phys}}}} \pi\left(\mathcal{R}^* = \mathcal{R}\middle|\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}\right)\, \mathrm{d}\mathcal{R}^*\right] \pi\left(\boldsymbol{\theta}_{\mathrm{pi}}\middle|\boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right)}{\pi\left(\boldsymbol{\mathcal{R}}^*\middle|\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right)},$$

Physical + inadequacy models:

$$\pi\left(\boldsymbol{\theta}_{\mathrm{pi}}\middle|\mathcal{R}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) = \frac{\pi\left(\mathcal{R}\middle|\boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\pi\left(\boldsymbol{\theta}_{\mathrm{pi}}\middle|\boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{R}\middle|\boldsymbol{M}_{\mathrm{pi}}\right)},$$

Physical model + inadequacy model + noise model:

$$\begin{aligned}
&\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \mid \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) \\
&= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}}\, ;\, \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\, \pi\left(\boldsymbol{\theta}_{\mathrm{pi}}\middle|\boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}\middle|\boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\
&= \frac{\int_{\Omega_{\mathcal{R}}} \pi\left(\mathcal{R}^*\middle|\boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\, \pi\left(\mathcal{R}^*\middle|\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\, \mathrm{d}\mathcal{R}^*\, \pi\left(\boldsymbol{\theta}_{\mathrm{pi}}\middle|\boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}\middle|\boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)},
\end{aligned}$$

Multilevel Bayesian modeling of data in the presence of model inadequacy and noise

Consensus total error model?

$$\sigma^2_{\mathrm{total}} = \sigma^2_{\mathrm{inad}} + \sigma^2_{\mathrm{noise}}$$

$$\mathcal{N}\left(\mathcal{R}_i \middle| \mu_{\mathrm{inad}} + \mu_{\mathrm{noise}}, \sigma_{\mathrm{total}}\right) = \mathcal{N}\left(\mathcal{R}_i \middle| \mu_{\mathrm{inad}}, \sigma_{\mathrm{inad}}\right) \otimes \mathcal{N}\left(\mathcal{R}_i \middle| \mu_{\mathrm{noise}}, \sigma_{\mathrm{noise}}\right)$$

Physical + noise models:

$$\pi\left(\boldsymbol{\theta}_{\mathrm{phys}} \,\middle|\, \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right) = \frac{\left[\int_{\boldsymbol{\mathcal{R}}^*{}_{\theta_{\mathrm{phys}}}} \pi\left(\mathcal{R}^* = \mathcal{R} \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}\right)\, \mathrm{d}\mathcal{R}^*\right] \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right)}{\pi\left(\boldsymbol{\mathcal{R}}^* \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{phys}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{phys}}}\right)},$$

Physical + inadequacy models:

$$\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \mathcal{R}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right) = \frac{\pi\left(\mathcal{R} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right) \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{R} \middle| \boldsymbol{M}_{\mathrm{pi}}\right)},$$

Physical model + inadequacy model + noise model:

$$\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \,\middle|\, \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)$$

$$\mathcal{N}\left(\mathcal{R}^* \middle| \mu_{\mathrm{inad}}, \sigma_{\mathrm{inad}}\right)$$

$$= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}}\ ;\ \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}$$

$$\mathcal{N}\left(\mathcal{R}^* \middle| \mu_{\mathrm{noise}}, \sigma_{\mathrm{noise}}\right)$$

$$= \frac{\int_{\boldsymbol{\Omega}_{\mathcal{R}}} \pi\left(\mathcal{R}^* \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\ \pi\left(\mathcal{R}^* \middle| \mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\ \mathrm{d}\mathcal{R}^*\ \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \Theta_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)},$$

Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

I_{obs} : observed fluorescence intensity [RFU]

I_{tru} : tumor cells intensity [RFU]

I_{bac} : background intensity [RFU]

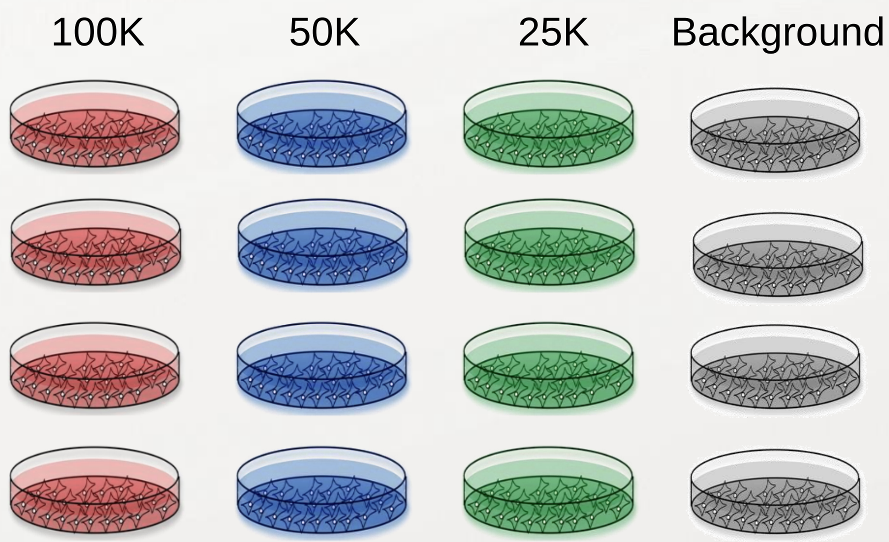
$$I_{\mathrm{obs}} = I_{\mathrm{tru}} + I_{\mathrm{bac}}$$

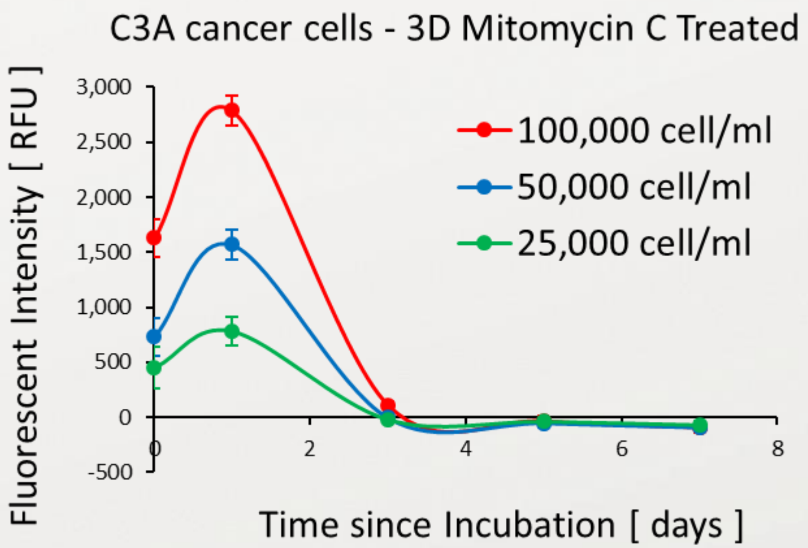
$$\Rightarrow \quad \widehat{I}_{\mathrm{tru}} = I_{\mathrm{obs}} - \widehat{I}_{\mathrm{bac}}$$

$$\leq \mathbf{0}$$

$\widehat{I}_{\mathrm{tru}}$ [RFU]

day 7				
concentration	**100K**	**50K**	**25K**	**1σ error**
sample 1	**-46.3**	**-126**	**-59.3**	±7.1
sample 2	**-59.3**	**-41.3**	**-58.3**	±7.1
sample 3	**-157**	**-119**	**-114**	±7.1
sample 4	**-96.3**	**-70.3**	**-45.3**	±7.1



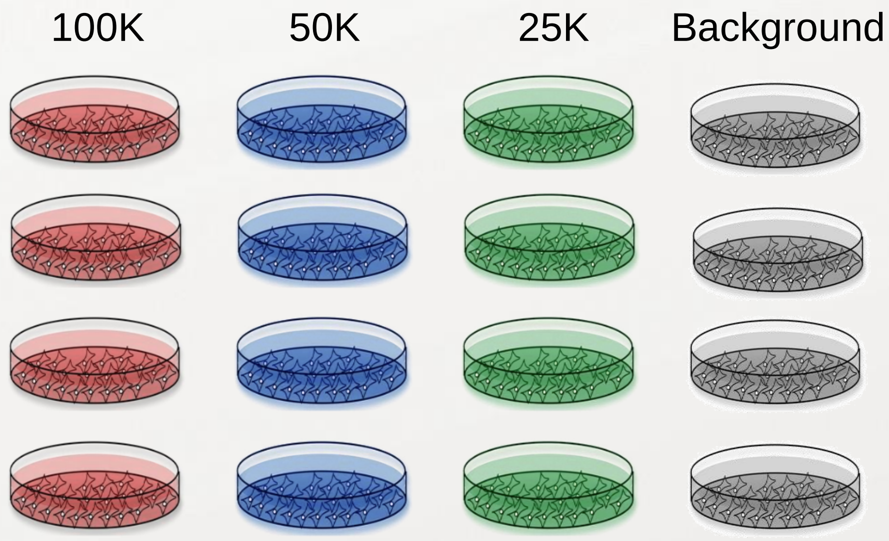


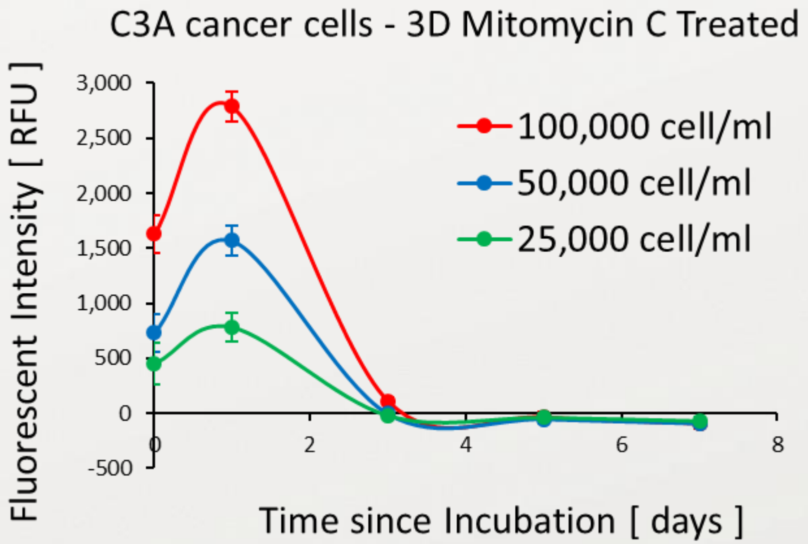
Traditional orthodox solutions can lead to logical paradoxes.
How to infer data from data?!

Viability of C3A immortalized liver tumor cells treated with Mitomycin-C (**MC**)

Physical model + inadequacy model + noise model:

$$
\begin{aligned}
\pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \;\middle|\; \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)
&= \frac{\mathcal{L}\left(\boldsymbol{\theta}_{\mathrm{pi}} \;;\; \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\, \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} \\
&= \frac{\displaystyle\int_{\boldsymbol{\Omega}_{\mathcal{R}}} \pi\left(\mathcal{R}^{*} \middle| \boldsymbol{\theta}_{\mathrm{pi}}, \boldsymbol{M}_{\mathrm{pi}}\right)\, \pi\left(\mathcal{R}^{*} \middle| \mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}}, \mathcal{I}_{\mathcal{R}}\right)\, \mathrm{d}\mathcal{R}^{*}\; \pi\left(\boldsymbol{\theta}_{\mathrm{pi}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)}{\pi\left(\mathcal{D}, \mathit{\Theta}_{\mathrm{nois}}, \mathcal{M}_{\mathrm{nois}} \middle| \boldsymbol{M}_{\mathrm{pi}}, \mathcal{I}_{\mathcal{R}}, \mathcal{I}_{\boldsymbol{\theta}_{\mathrm{pi}}}\right)} ,
\end{aligned}
$$





References

1. Amir Shahmoradi, 2017, Multilevel Bayesian Parameter Estimation in the Presence of Model Inadequacy and Data Uncertainty, arXiv:1711.10599 [physics.data-an]

```
@article{shahmoradi2017multilevel,
  title={Multilevel Bayesian Parameter Estimation in the Presence of Model Inadequacy and Data Uncertainty},
  author={Shahmoradi, Amir},
  journal={arXiv preprint arXiv:1711.10599},
  year={2017}
}
```