Optimization and Simulation

Introduction

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

Course map

Model & Uncertainty

Specify inputs, randomness, and mechanisms

Simulation

Generate synthetic outcomes and assess performance

Decide / Optimize

Search for better configurations under uncertainty

Course map

Lecture sequence

1. Introduction to simulation
2. Drawing from distributions
3. The Poisson process
4. Discrete events simulation
5. Statistical analysis and bootstrapping
6. Variance reduction
7. Markov chain Monte Carlo methods
8. Introduction to optimization (heuristics)
9. Multi-objective optimization

Course map

What you will be able to do

- Build a simulator and generate reliable outputs (not just one run).
- Quantify uncertainty (confidence intervals, bootstrap) and reduce variance.
- Use Monte Carlo and MCMC to sample complex distributions.
- Optimize decisions using simulation outputs (single- and multi-objective).

Outline

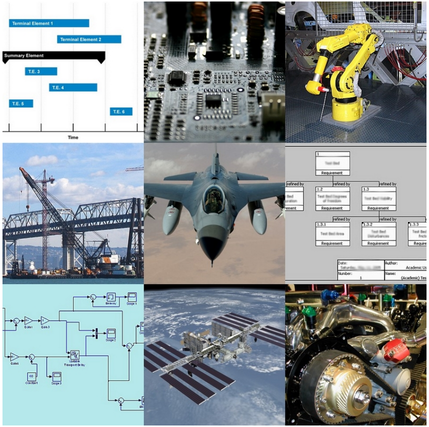
Engineering systems

Definition (Wikipedia)

Combination of components that work in synergy to collectively perform a useful function.

Properties

- Complex
- Large
- Designed
- Configurable
- Interactions with external world



Source: Wikipedia

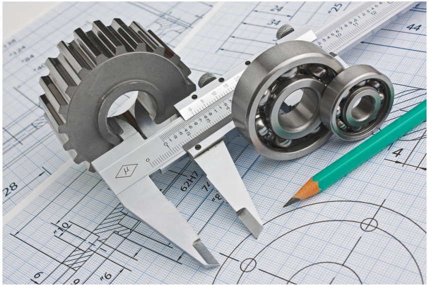
Engineering systems

Objectives

- Design
- Maintain
- Operate

Time horizon

- Long-term
- Medium-term
- Short-term

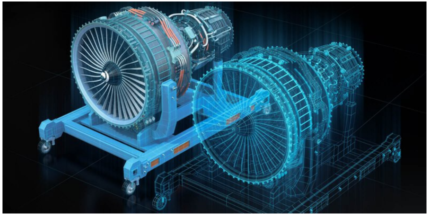


Source: Swiss Learning Exchange

Engineering systems

Mathematical and digital twins

- Modeling
- Simulation
- Optimization



Source: Konica Minolta

Engineering systems

	Modeling	Simulation	Optimization
Roles	Represent	Predict	Improve
How?	Capture causal effects	Capture the propagation of uncertainty	Investigate better configurations

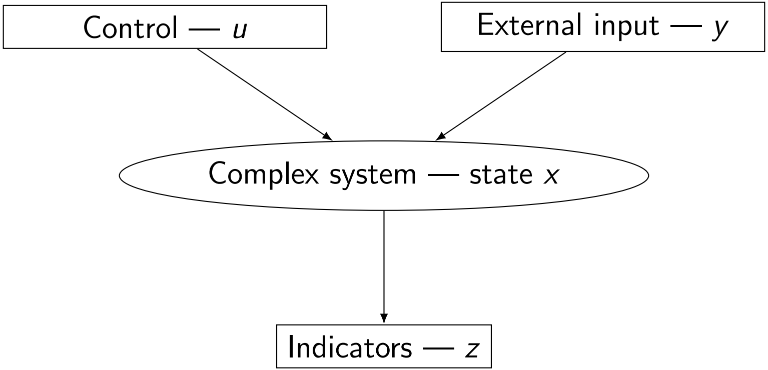
Outline

Modeling

System

A system can be represented as follows:

$$z = h(x, y, u; \theta)$$



Modeling

$$z = h(x, y, u; \theta)$$

Example

A car:

- x captures the state of the system (e.g. speed, position of other vehicles)
- y captures external influences (e.g. wind)
- u captures possible human controls on the system (e.g. acceleration/deceleration)
- z represents indicators of performance (e.g. oil consumption).

Modeling

Decompose the complexity

- The model h is usually decomposed to reflect the interactions of the subsystems
- For example,
  - a car-following model captures the target speed of the driver,
  - an engine model derives the actual consumption as a function of the acceleration.

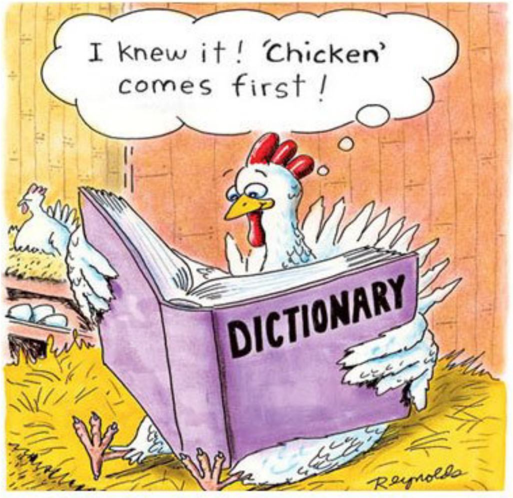
Modeling

Causal effects

- Very important to identify the causal effects
- Failure to do so may generate wrong conclusions

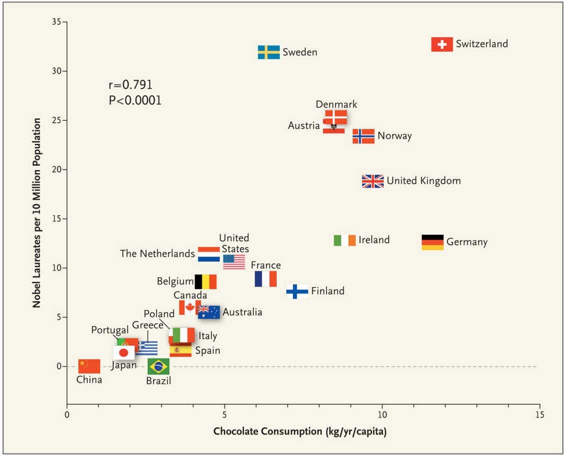
Forecasting

Assumption: causal effects are stable over time and configurations of the system.



Data can be misleading

Chocolate Consumption, Cognitive Function, and Nobel Laureates



Source: [Messerli, 2012]

Inference

Data collection

On an existing system, collect N observations of x_n, y_n, u_n, z_n, $n = 1, \ldots, N$.

Goodness of fit

For a given value of θ, "distance" $d_n(\theta)$ between

- the predicted value $h(x_n, y_n, u_n; \theta)$, and
- the observed value z_n.

Inference

Find $\widehat{\theta}$ that minimizes the total distance:

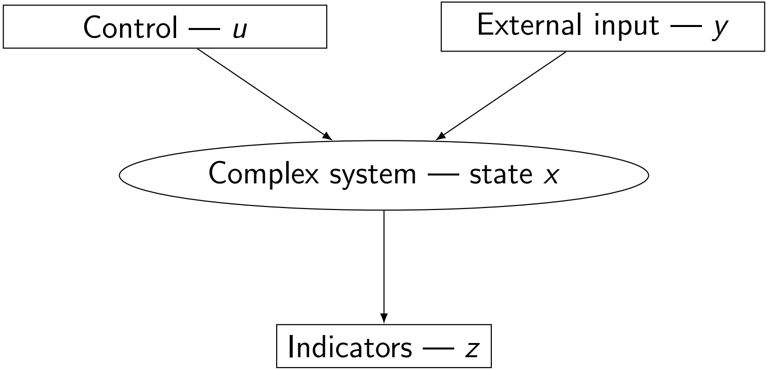
$$\widehat{\theta} = \mathrm{argmin}_{\theta} \sum_{n=1}^{N} d_n(\theta).$$

Outline

Simulation

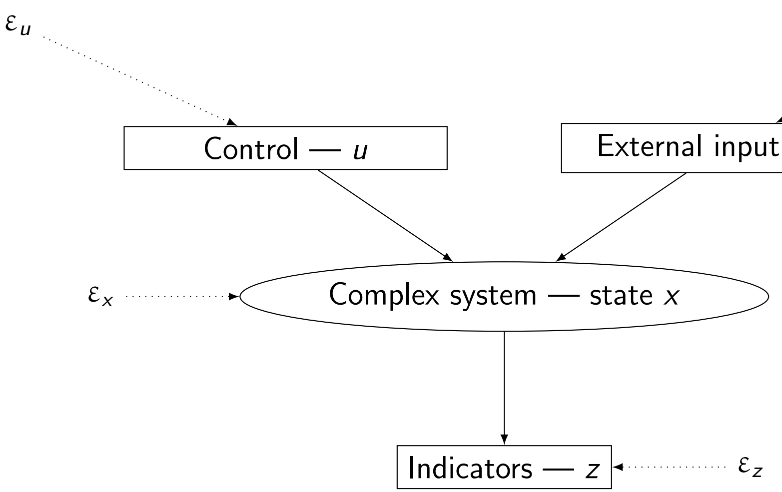
Simulation is more than simply applying the model.

$$z = h(x, y, u; \widehat{\theta})$$



Simulation

$$Z = h(X, Y, U; \widehat{\theta}) + \varepsilon_z$$



Simulation

Propagation of uncertainty

$$Z = h(X, Y, U; \widehat{\theta}) + \varepsilon_z$$

- Given the distribution of X, Y, U and ε_z
- what is the distribution of Z?

Why is it difficult?

Arithmetic of random variables

- Let X and Y be two random variables (assume independent for now)
- Define a new random variable

$$Z = X + Y$$

- What is the probability distribution of Z?

Analytical solution

The probability density function of Z is given by a convolution:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z - x)\, dx$$

Independence is convenient, dependence is common

In practice, variables are rarely independent

- Correlated inputs: demand, weather, incidents, prices.
- Common shocks: one event affects many components simultaneously.
- **Time dependence:** queues, inventories, fatigue, learning.

Takeaway

- Convolutions become harder; analytical results become rarer.
- Simulation remains straightforward: sample the joint distribution and propagate it through h.

Why propagate uncertainty?

Key difficulty

- Convolution integrals quickly become intricate.
- Closed-form expressions are rare.
- The situation worsens for nonlinear functions or many variables.
- Independence is convenient, but dependence is common.
- Convolutions become harder; analytical results become rarer.

Example: sum of two uniform random variables

Assumptions

$$X \sim \mathrm{U}(0,1), \qquad Y \sim \mathrm{U}(0,1), \qquad Z = X + Y$$

Result (after convolution)

The density of Z is triangular:

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ z, & 0 \leqslant z \leqslant 1, \\ 2 - z, & 1 < z \leqslant 2, \\ 0, & z > 2. \end{cases}$$

Simulation: propagate uncertainty numerically

Core idea

- Generate numerical values that follow the distribution of interest
- Apply standard arithmetic to these values
- Let empirical distributions replace analytical formulas

Draws from a random variable

- Let X have pdf f_X
- A draw from X is a numerical value generated according to f_X
- Consider independent draws $X_1, \ldots, X_R$

Simulation: propagate uncertainty numerically

Key convergence property (Law of Large Numbers)

Assume $X_1, \ldots, X_R$ are **i.i.d.** draws from f_X. For any interval $[a, b]$,

$$\frac{1}{R}\sum_{i=1}^{R} \mathbf{1}\{X_i \in [a, b]\} \xrightarrow[R\to\infty]{\text{a.s.}} \int_a^b f_X(x)\, dx.$$

Preview

Later, MCMC generates **dependent** draws; convergence still holds under additional conditions.

Illustration: simulation replaces convolution

Procedure

- Draw $X_1, \ldots, X_R \sim \mathrm{U}(0, 1)$
- Draw $Y_1, \ldots, Y_R \sim \mathrm{U}(0, 1)$
- Compute
$$Z_r = X_r + Y_r, \qquad r = 1, \ldots, R$$
- Analyze the empirical distribution of Z_r

Illustration: simulation replaces convolution

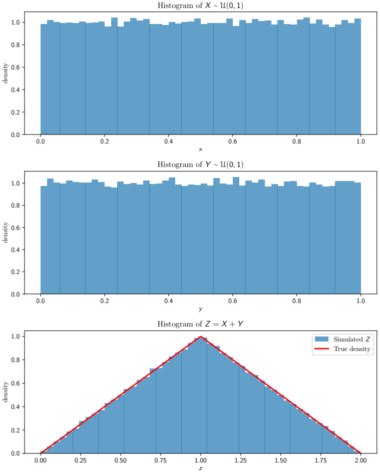


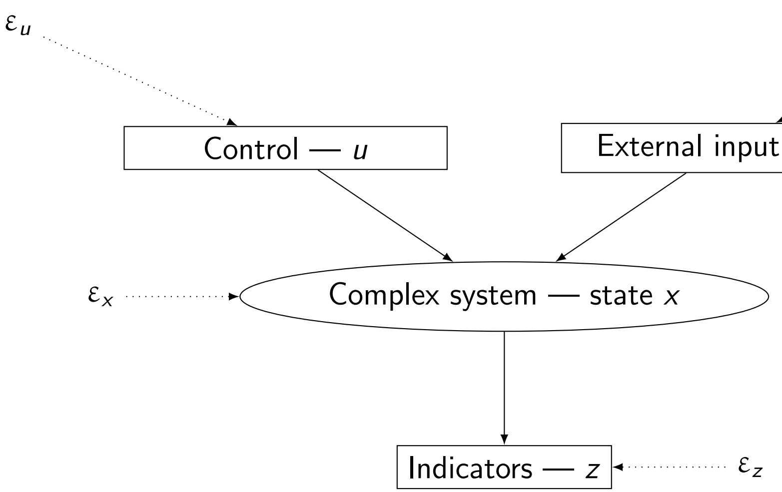
Illustration: simulation replaces convolution

Key message

- No integrals, no convolutions
- Only draws and simple arithmetic
- The distribution of Z emerges automatically

Simulation

$$Z = h(X, Y, U; \widehat{\theta}) + \varepsilon_z$$



Simulation

Sampling

- Draw realizations of X, Y, U, ε_z
- Call them $x^r, y^r, u^r, \varepsilon_z^r$
- For each r, compute

$$z^r = h(x^r, y^r, u^r; \widehat{\theta}) + \varepsilon_z^r$$

- z^r are draws from the random variable Z

Analysis

- Generate many draws from Z.
- Analyze their empirical distribution.

Importance of number of draws

Theory vs. practice

- Theory: true distribution of Z when $r \to \infty$.
- Practice: finite number R of draws.
- If R is too small, simulator output is just noise.

Analogy with real world

- Nature also generates instances of a complex random variable.
- Experiments must be repeated in order to reach conclusions.

Example: policy analysis

- The real impact of a policy is difficult to analyze.
- Incomplete results consistent with expectations may lead to erroneous conclusions.

Two sources of variability

Two different questions

- **Output uncertainty (system variability):** how variable is Z itself?
- **Monte Carlo error (estimation noise):** how accurate is $\bar{Z}_R$ as an estimate of $\mathbb{E}[Z]$?

Key distinction

- Even if Z is highly variable, $\bar{Z}_R$ can be estimated accurately with large R.
- Even if Z is not very variable, $\bar{Z}_R$ can look noisy when R is small.

Typical relationship (i.i.d. runs)

$$\mathrm{Var}(\bar{Z}_R) = \frac{\mathrm{Var}(Z)}{R} \qquad \Rightarrow \qquad \text{Monte Carlo error decreases as } 1/\sqrt{R}.$$

Example: enhancing road safety

Traffic accidents in Switzerland

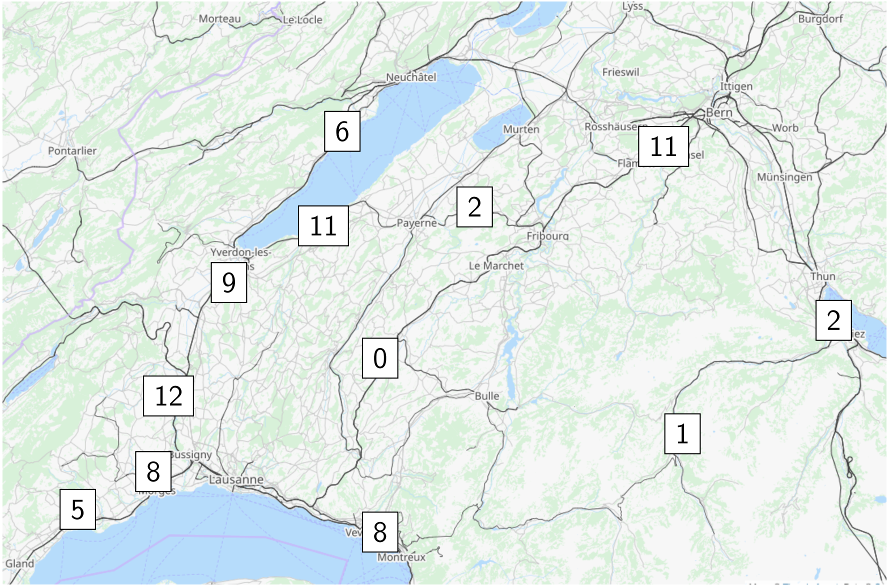
- The Swiss Federal Roads Office has proposed new safety measures.
- The initiative aims to identify accident-prone locations and significantly lower speed limits in those areas.
- The TCS has organized a referendum opposing these measures, arguing that they are not sufficiently effective in improving safety and significantly hinder mobility.
- You have been tasked with evaluating the actual effectiveness of these safety measures.

Example: enhancing road safety

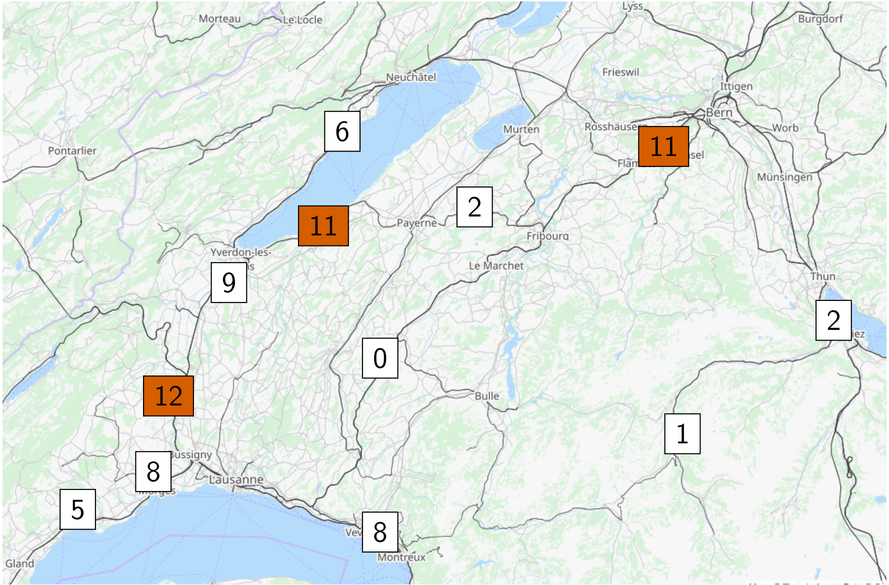
Methodology: simulation

- Simulate the number of accidents before and after the implementation of the safety measures.
- Based on existing literature, the number of accidents in Switzerland can be modeled using a uniform distribution ranging from 0 to 12.

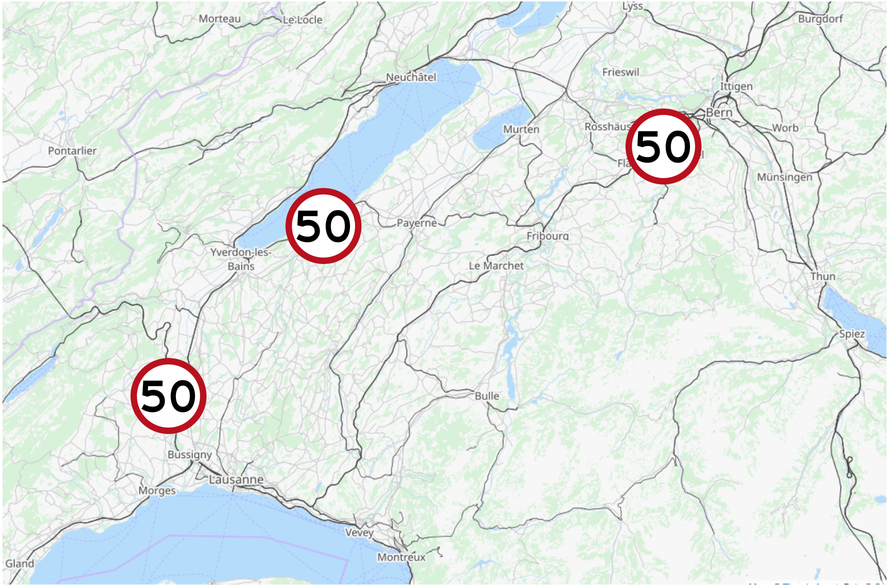
Current situation: simulated number of accidents



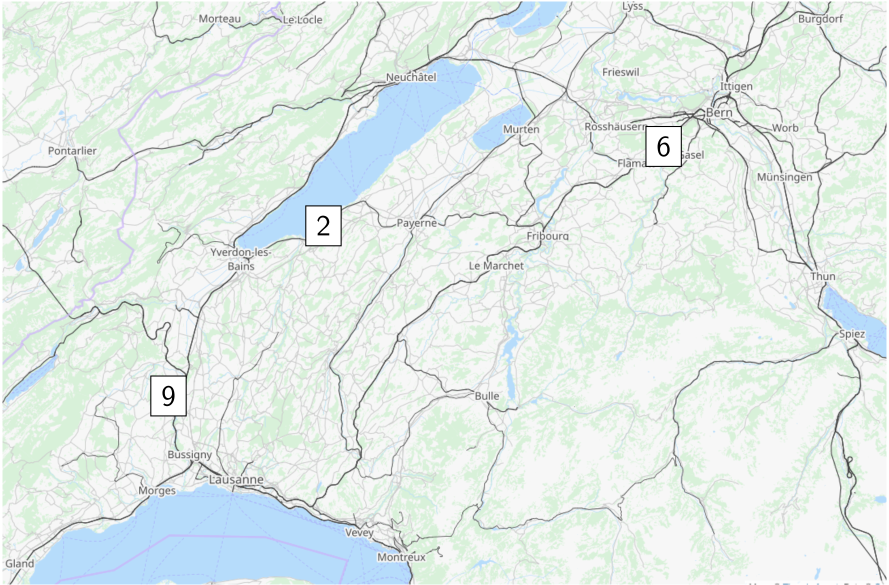
Identification of dangerous spots



Implementation of the measures



After implementation: simulated number of accidents



Analysis

Data

- Before: $12 + 11 + 11 = 34$ accidents
- After: $9 + 2 + 6 = 17$ accidents
- 50% reduction.

Conclusion

As expected, safety measures reduce the number of accidents.

Example: enhancing road safety

Major flaws

- Causal effects are not accounted for in the model.
- The simulation was performed using only a single run.
- Not all locations were re-simulated.
- Confidence in the conclusions is strengthened by their alignment with expectations.

What should have been done

- Run multiple simulations to estimate accident occurrences more robustly.
- The expected average number of accidents should be around 6 at all locations, regardless of the speed limit.
- A formal statistical test would likely fail to reject the null hypothesis, suggesting that the speed limit has no measurable effect in the simulation.

Outline

Simulation

Derivation of indicators from the distribution

- Mean
- Variance
- Modes
- Quantiles

Statistics

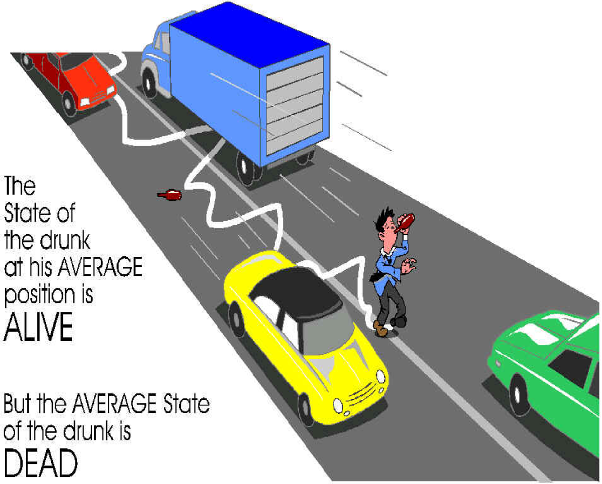
"Numbers don't lie. That's where we come in."

Indicators

- Mean: $\mathrm{E}[Z] \approx \bar{Z}_R = \frac{1}{R} \sum_{r=1}^{R} z^r$
- Sample variance:
  $\mathrm{Var}(Z) \approx s_R^2 = \frac{1}{R-1} \sum_{r=1}^{R} (z^r - \bar{Z}_R)^2$
- Modes: based on the histogram
- Quantiles: sort and select

Important: there is more than the mean

The mean



[Savage et al., 2012]

The mean

The flaw of averages

[Savage et al., 2012]

$$\mathrm{E}[Z] = \mathrm{E}[h(X, Y, U; \widehat{\theta}) + \varepsilon_z] \neq h(\mathrm{E}[X], \mathrm{E}[Y], \mathrm{E}[U]; \widehat{\theta}) + \mathrm{E}[\varepsilon_z]$$

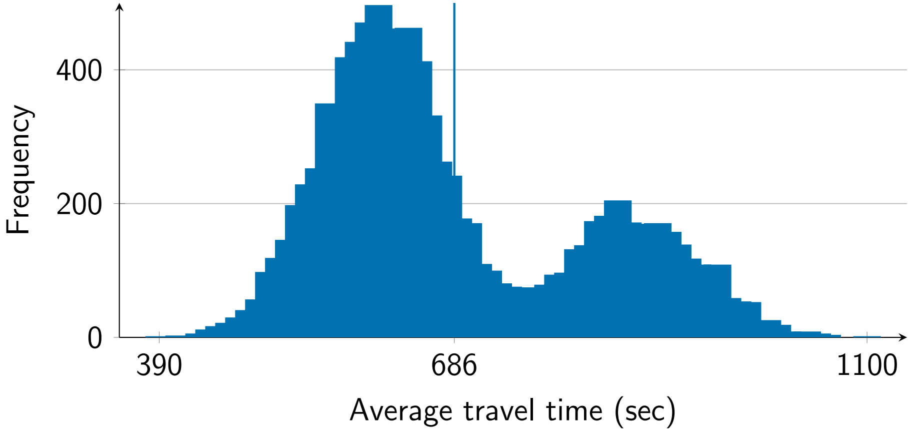
... except if h is linear.

There is more than the mean

Example

- Intersection with capacity 2000 veh/hour
- Traffic light: 30 sec green / 30 sec red
- Constant arrival rate: 2000 veh/hour during 30 minutes
- With 30% probability, capacity at 80%.
- Indicator: Average time spent by travelers

There is more than the mean



Pitfalls of simulation

Few number of runs

- Run time is prohibitive
- Tempting to generate partial results rather than no result

Focus solely on the mean

- The mean is useful, but not sufficient.
- For complex distributions, it may be misleading.
- Intuition from normal distribution (mode = mean, symmetry) do not hold in general.
- Important to investigate the whole distribution.
- Simulation allows to do it easily.

Challenges

- How to generate draws from Z?
- How to represent complex systems? (specification of h)
- How large R should be?
- How good is the approximation?

Pseudo-random numbers

Definition

- Deterministic sequence of numbers
- which have the appearance of draws from a $U(0,1)$ distribution

Typical sequence

$$x_n = ax_{n-1} \text{ modulo } m$$

- This has a period of the order of m
- So, m should be a large prime number
- For instance: $m = 2^{31} - 1$ and $a = 7^5$
- x_n/m lies in the $[0,1[$ interval

Pseudo-random numbers

Modern practice

- We almost never implement RNGs ourselves.
- We use tested generators from standard libraries (period, equidistribution, statistical quality).
- What matters in practice: seeds (reproducibility) and independent replications.

Pseudo-random numbers

Python/Numpy example (reproducibility)

```
import numpy as np

rng1 = np.random.default_rng(2026)
x1 = rng1.uniform(0.0, 1.0, size=5)
print(x1)

rng2 = np.random.default_rng(2026)
x2 = rng2.uniform(0.0, 1.0, size=5)
print(x2)
[0.17893481 0.63991317 0.4672684  0.37050053 0.35491733]
[0.17893481 0.63991317 0.4672684  0.37050053 0.35491733]
```

Pseudo-random numbers

Python/Numpy example (independent replications)

```
import numpy as np

master = np.random.default_rng(12345)
seeds = master.integers(0, 2**63 - 1, size=3)

for r, s in enumerate(seeds):
    rng = np.random.default_rng(int(s))
    x = rng.normal(size=1000)
    print(f"Replication␣{r}:␣mean={x.mean():.3f},␣std={x.std()
```

```
Replication 0: mean=0.014, std=1.009
Replication 1: mean=-0.015, std=0.992
Replication 2: mean=-0.058, std=0.999
```

Pseudo-random numbers

What this shows

- Different replications produce different realizations.
- Summary statistics fluctuate around their theoretical values.
- Re-running the script produces identical output (reproducibility).

Outline of the lectures on simulation

- Drawing from distributions
- Discrete event simulation
- Data analysis
- Variance reduction
- Markov Chain Monte Carlo

Reference

[Ross, 2012]

Outline

Optimization

Assumptions

- Control U is deterministic.

$$Z(u) = h(X, Y, u) + \varepsilon_z$$

- Various features of Z are considered: mean, variance, quantile, etc.

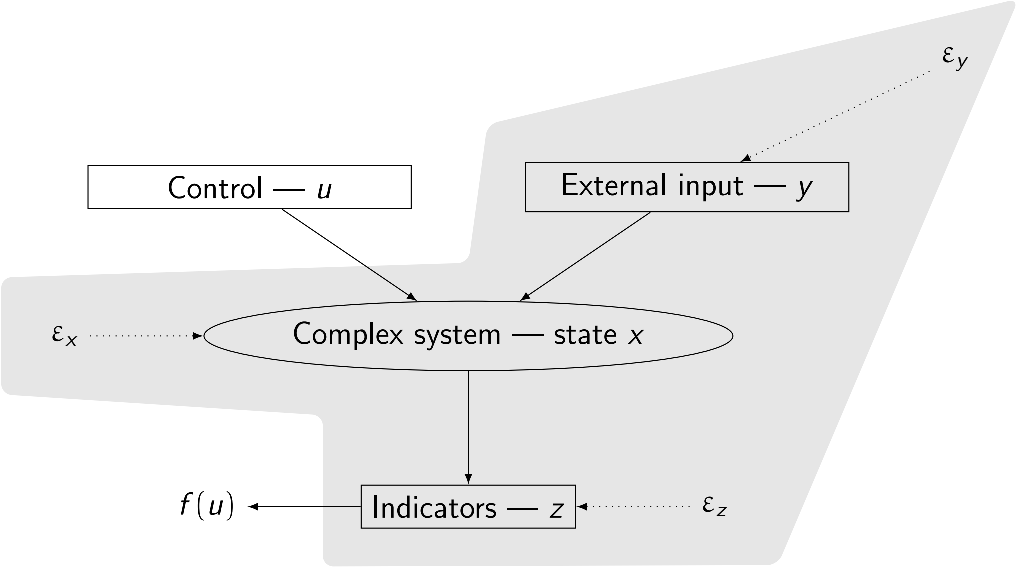
$$(z_1(u), \ldots, z_m(u))$$

- They are combined in a single indicator:

$$f(u) = g(z_1(u), \ldots, z_m(u))$$

- If not, it is called multi-objective optimization.

General framework: the black box



Optimization problem

$$\min_{u \in \mathbb{R}^n} f(u)$$

subject to

$$u \in \mathcal{U} \subseteq \mathbb{R}^n$$

- u: decision variables
- $f(u)$: objective function
- $u \in \mathcal{U}$: constraints
- $\mathcal{U}$: feasible set

In this course...

- Classical optimization problems
- Heuristics
- Multi-objective optimization

Summary

Modeling

- Decomposition of the complexity.
- Causal effects.

Simulation

- Propagation of uncertainty.
- Requires many draws.
- Analysis of the entire empirical distribution.
- There is more than the mean.

Optimization

- Identify the control that improves a function of the indicators.
- Optional: multi-objective optimization.

Bibliography

Messerli, F. (2012).
Chocolate consumption, cognitive function, and Nobel laureates.
The New England Journal of Medicine, 367:1562–1564.

Ross, S. (2012).
Simulation.
Academic Press, fifth edition edition.

Savage, S., Danziger, J., and Markowitz, H. (2012).
The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty.
Wiley.