Stochastic Simulation

Cathy Wu

1.041/1.200 Transportation: Foundations and Methods

Readings

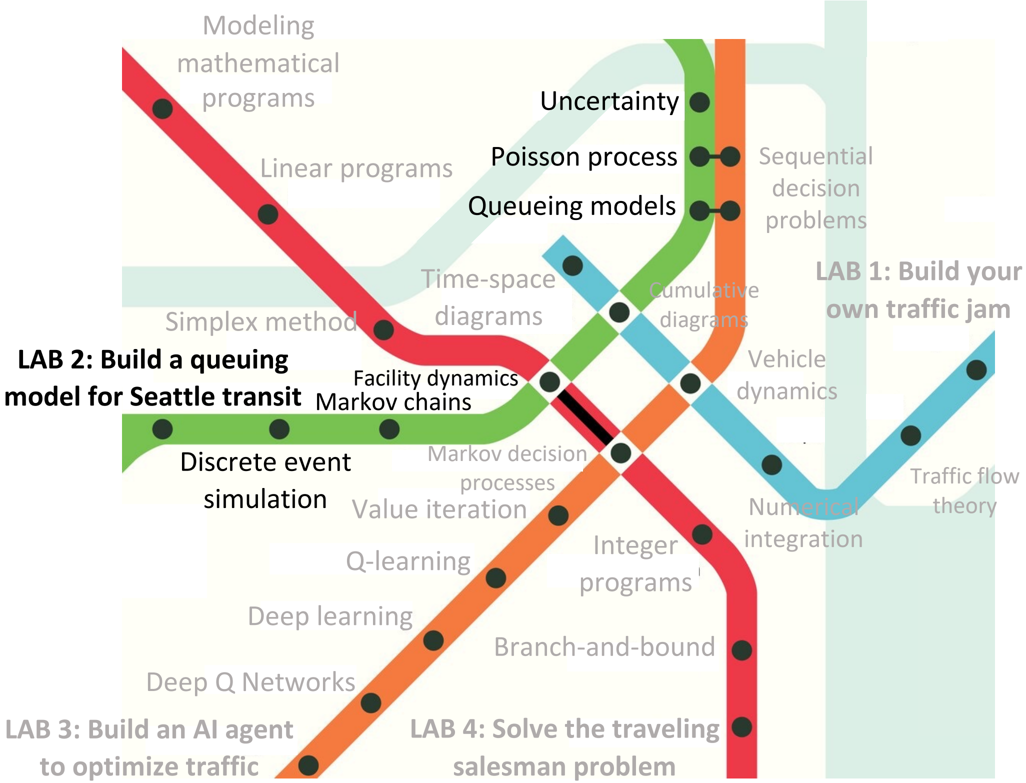
1. Larson, Richard C. and Amedeo R. Odoni. **Urban Operations Research**. Prentice-Hall (1981). Chapter 7: Simulations. URL.
2. (Optional) X. Wang *et al.*, “Traffic light optimization with low penetration rate vehicle trajectory data,” *Nat Commun*, vol. 15, no. 1, Art. no. 1, Feb. 2024, doi: 10.1038/s41467-024-45427-4.

Unit 2: Queuing systems

Unit 2

Modeling

Stochastic



Traffic light optimization with low penetration rate vehicle trajectory data (Nature Communication, 2024)

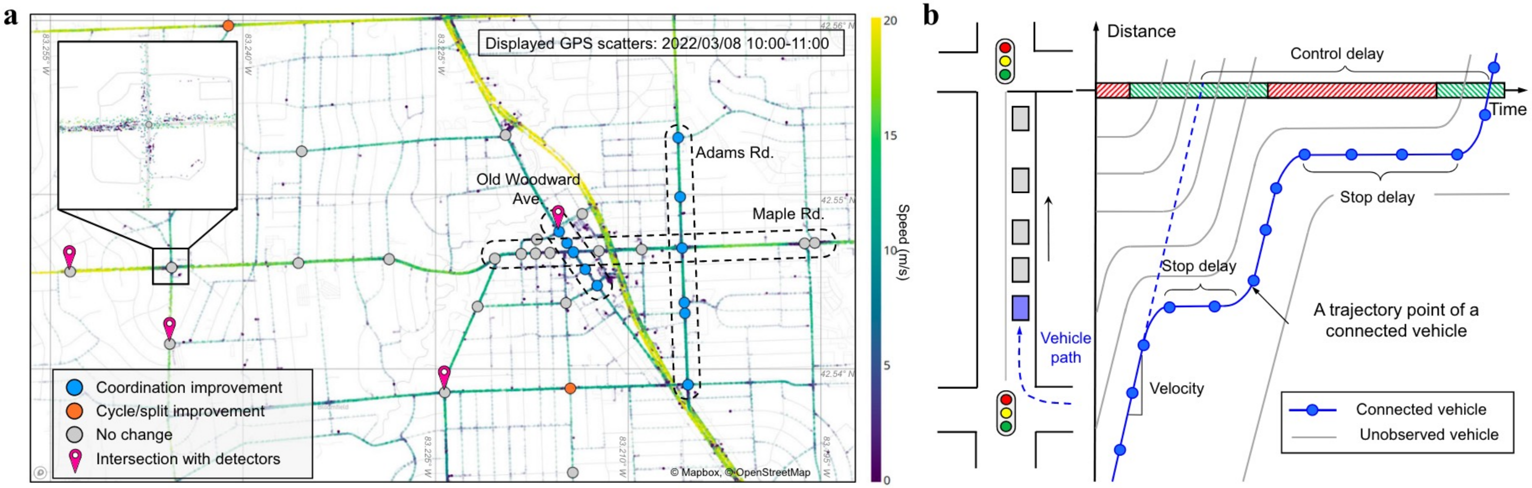


Fig. 1 | Traffic signal retiming with vehicle trajectories.

- Overall result: Decreased the delay and number of stops at signalized intersections by up to 20% and 30% with data from as low as 6% vehicle penetration.

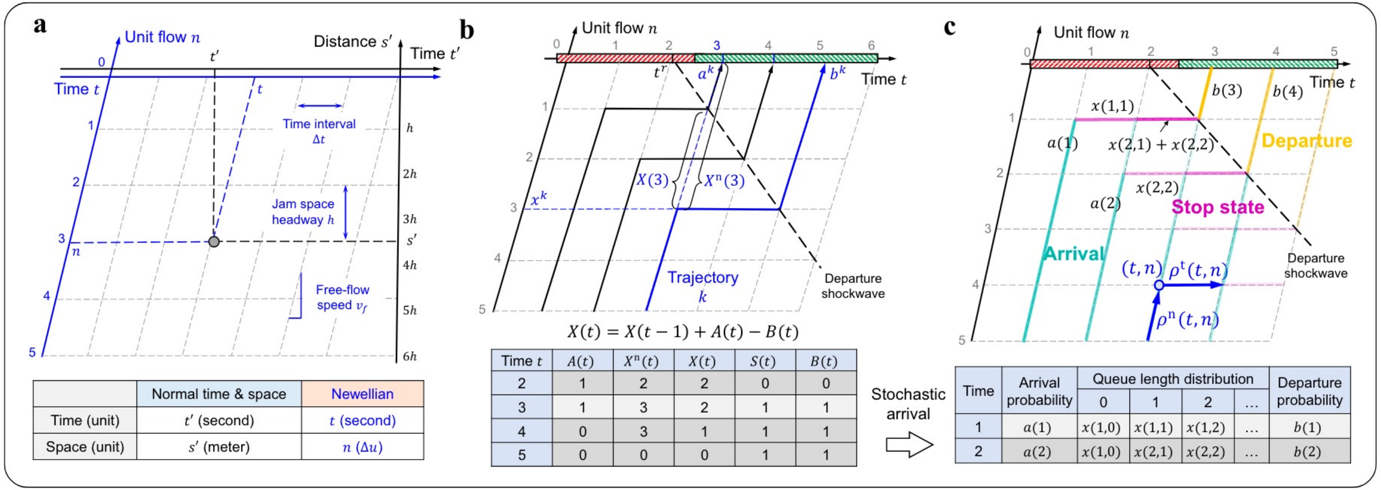
X. Wang *et al.*, "Traffic light optimization with low penetration rate vehicle trajectory data," *Nat Commun*, vol. 15, no. 1, Art. no. 1, Feb. 2024, doi: 10.1038/s41467-024-45427-4.

Ideas from Unit 1+2

- Time-space diagram
- Newell's car following model
- Stochastic arrival process
  - Bernoulli distribution!

- Discrete event simulation
- G/D/1 queue

→ Probabilistic time-space (PTS) diagram



	Normal time & space	Newellian
Time (unit)	t' (second)	t (second)
Space (unit)	s' (meter)	n (Δu)

Time t	$A(t)$	$X^n(t)$	$X(t)$	$S(t)$	$B(t)$
2	1	2	2	0	0
3	1	3	2	1	1
4	0	3	1	1	1
5	0	0	0	1	1

Time	Arrival probability	Queue length distribution				Departure probability
		0	1	2	…	
1	$a(1)$	$x(1,0)$	$x(1,1)$	$x(1,2)$	…	$b(1)$
2	$a(2)$	$x(1,0)$	$x(2,1)$	$x(2,2)$	…	$b(2)$

Fig. 2 | Point-queue model under Newellian coordinates and PTS diagram.

X. Wang *et al.*, "Traffic light optimization with low penetration rate vehicle trajectory data," *Nat Commun*, vol. 15, no. 1, Art. no. 1, Feb. 2024, doi: 10.1038/s41467-024-45427-4.

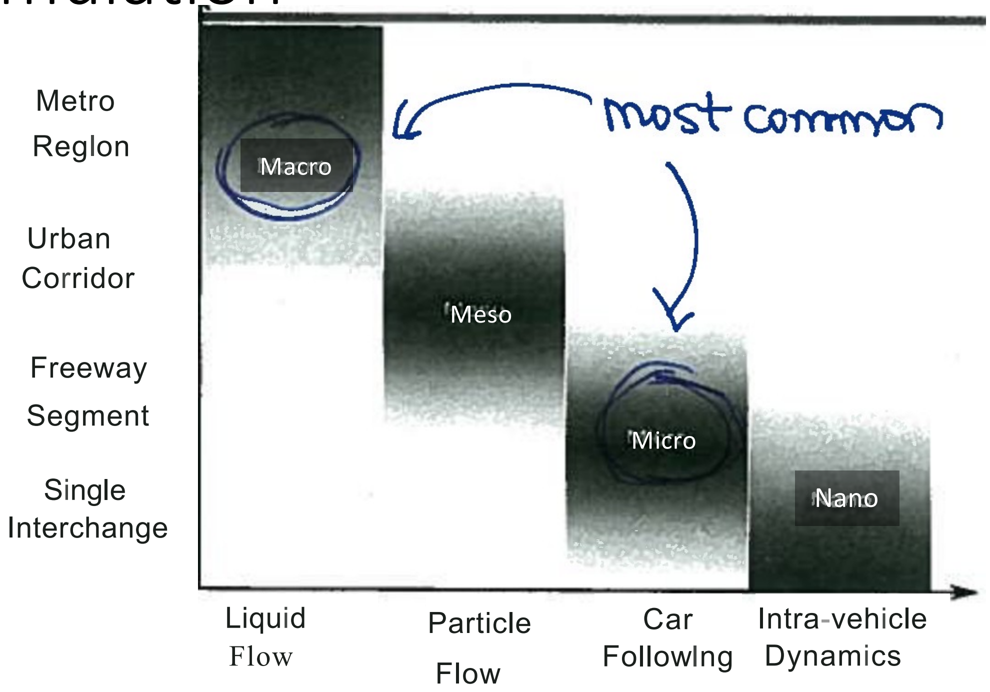
Outline

1. Introduction to simulation
2. Discrete event simulation
3. Transient analysis
4. Mixed continuous and discrete event simulation

Outline

1. **Introduction to simulation**
   a. Urban traffic simulation
   b. Simulation pro's and con’s
2. Discrete event simulation
3. Transient analysis
4. Mixed continuous and discrete event simulation

Traffic simulation



- Source: "Scale and Complexity Tradeoffs In Surface Transportation Modeling", Karl E. Wunderlich.

Stochastic simulation models

- Def. **Simulation**: Modeling approach, the aim of which is to approximate (i.e. simulate) the systems' behaviour with the help of models that:
  - are computer-based models that try to imitate the behavior of a physical system
  - account for uncertainty: computationally mimic randomness, i.e. simulate random events
- Why simulate?
  - Simple simulations enable further understanding the concepts of random variables, their distributions and realizations
  - Understanding the main underlying concepts of a simulation model enables: understanding the complexity and need for a rigorous validation of simulation models and statistical analysis of outputs

Simulation pro's and con’s

- Advantages
  - May be suitable for problems that are not analytically tractable
  - Greater level of modeling detail (does not necessarily mean increased realism)
  - Allows for simulated experiments of otherwise costly (e.g. high risk) or infeasible field experiments
  - What-if analysis: trial and error procedure
  - Useful to test the validity of mathematical assumptions (e.g. to validate analytical models)
- Disadvantages
  - What-if analysis: difficult to develop causal relationships
  - Computationally expensive mathematical tool (need for many replications)
  - Proper statistical analysis of the outputs is complex
  - Detailed model requires very detailed data to be formulated and calibrated
  - Data quality, "garbage-in, garbage-out"
  - Difficult to use to perform optimization (simulation-based optimization)
- Just as analytical models, simulation models are based on numerous assumptions and approximations, use it with caution and keep in mind that it's a simplification of reality, i.e. a MODEL!

Stochastic simulation models

- How to simulate a simple queue?
- How to advance time?
- Need to mimic randomness in a computer, i.e. simulate random events
  - How to generate random numbers?
  - How to generate random observations from a probability distribution?

Outline

1. Introduction to simulation
2. **Discrete event simulation**
   a. Simulation of an $M/M/1$ queueing system
   b. Random number generation
   c. Replications
3. Transient analysis
4. Mixed continuous and discrete event simulation

Simulation models for queuing systems

Discrete-event simulation (DES)

- Continuous-time: event driven simulation
- The system is modeled by a set of discrete states
- The system can change states when an **event** occurs
- Between events the system does not change states
- Identify the events that lead to state changes
- For each event, describe:
  - the new state
  - changes in the system attributes
  - triggered events
- The simulator keeps an *event list* of the events that are scheduled to happen with their scheduled time
  The events are ordered chronologically
- Once an event is carried out, the simulation dock is advanced to the time the next event is scheduled to start

Discrete event simulation model: M/M/1

- State of the system: $N(t)$ = number of customers in the queueing system at time t
- Events:
  - arrival of a new customer
  - service completion for the customer currently in service
- Update system state and record information at time of each event
- State transition mechanism for event-driven simulation:

$$N(t) = \begin{cases} N(\text{preceding event}) + 1, \text{if an arrival occurs at time } t \\ N(\text{preceding event}) - 1, \text{if a service completion occurs at time } t \end{cases}$$

- Simulation end: if the simulation clock exceeds a pre-specified (simulation) time or number of customers observed reaches a specified limit
- How to obtain the time and type of next event?

Discrete Event Simulation Model: M/M/1

- A_n: arrival time of customer n
- D_n: departure time (service completion time) of customer n
- H_n: inter-arrival (arrival headway) time of customer n
- S_n: service time of customer n

Random number generation

- Approach
  1. Generate "independent" draws from a uniform distribution,
  2. Then generate draws from an arbitrary distribution

Uniform random numbers

- A "random number" for computer simulation purposes is a random observation from the uniform distribution on the interval $[0, 1]$, i.e., sampled from $U[0,1]$

- Most software have functions that allow easy generation of random numbers.
  - Successive random numbers generated by `random.random()` can be considered mutually independent samples from $U[0,1]$
    - In Python: `import random; random.random()`
    - In Excel: `RAND()`
  - To return a matrix (n-by-m) of observations that can also be considered mutually independent samples from $U[0,1]$
    - In Python: `import numpy as np; np.random.rand(n, m)`
    - In Matlab: `rand(n,m)`

Uniform random numbers

- Pseudorandom number generator:
  - A deterministic algorithm for generating a sequence of numbers that approximates the properties of random numbers
- Most software use linear-congruential methods which involve modulo arithmetic Recursive algorithm:
$$r_{n+1} = (kr_n + a) \bmod m$$
where k, a, and m are positive integers ($k < m, a < m$).
Choose r_0, the seed
- What is a good random number generator?
  1. a long (maximum) period between repeat cycles
  2. covers unit interval uniformly
  3. same for higher dimensional unit-cubes
  4. does not show obvious patterns
  5. fools statistical tests when used in simulations
  6. efficiency of the algorithm

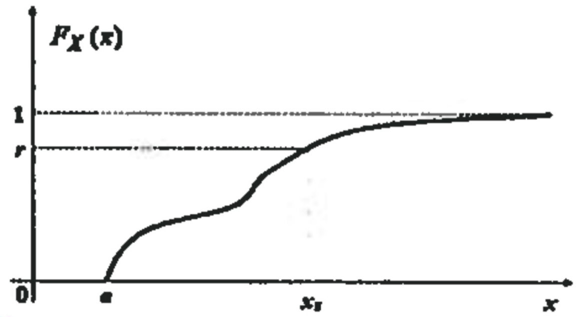
Discrete random number generation

- Based on $U(0,1)$ variables, generate variables from an arbitrary discrete dbn
- Discrete r.v. X has pmf:
$$P(X = x_k) = m_k, k = 1, \dots, K$$
- Split the unit interval into K subintervals of length $m_1, m_2, \dots, m_K$
- **Draw** a uniform random number r in $[0,1]$
- If r falls into the subinterval belonging to m_k then the r.v. realization is $X = x_k$
- Implementation:
    1. Let r be a draw from $U(0,1]$
    2. Initialize $k = 0, p = 0$
    3. $p = p + m_k$
    4. If $r < p$, set $X = x_k$ and stop.
    5. Otherwise, set $k = k + 1$ and go to step 3.

Continuous random number generation

- Most popular method: Inverse transform method (a.k.a. inverse function method)
  - Draw a uniform number r in $[0,1]$
  - Set $x = F_X^{-1}(r)$
- Based on the following property:
  - Let X be a continuous r.v. with cdf $F_X(x)$
  - Let U be a r.v. uniformly distributed over $[0,1]$
  - Then a new r.v. $\mathrm{Y} = F_X^{-1}(U)$ has F_X as its CDF
  - $P(F_X^{-1}(U) \leq r) = F_X(r)$
- Note: the cdf is not always available in closed-form: e.g. normal dbn

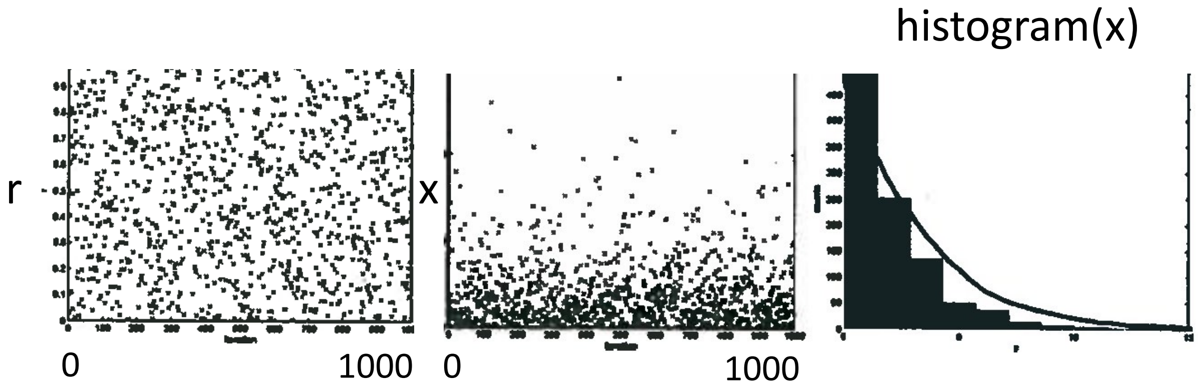
$$x_s = F_X^{-1}(r)$$



Example: exponential r.v.

```
sampleSize =1000;
lambda = 0.5;

r = rand(sampleSize,1);
x = -log(r)/lambda;
```



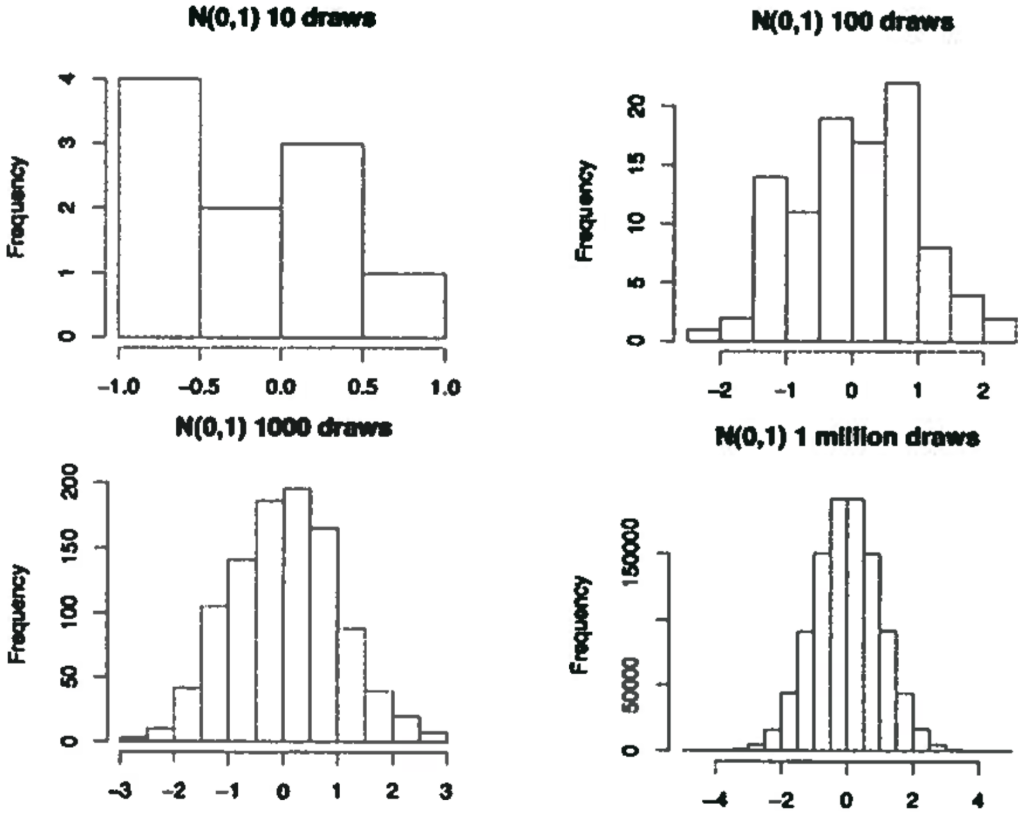
Discrete Event Simulation Model: M/M/1

- A_n: arrival time of customer n
- D_n: departure time (service completion time) of customer n
- H_n: inter-arrival (arrival headway) time of customer n
- S_n: service time of customer n
- $\begin{cases} A_1 = 0 \\ A_n = A_{n-1} + H_n, n = 2, 3, \ldots \end{cases}$
- $\begin{cases} D_1 = S_1 \\ D_n = S_n + \max\{A_n, D_{n-1}\}, n = 2, 3, \ldots \end{cases}$

  where $\max\{A_n, D_{n-1}\}$ represents the time that service starts for customer n
- H_n and S_n are generated from two exponential random variables. How can we generate these?

$$\begin{cases} H_n = -\dfrac{\ln(r_H)}{\lambda} = -\dfrac{\ln(rand())}{\lambda} \\ S_n = -\dfrac{\ln(r_S)}{\mu} = -\dfrac{\ln(rand())}{\mu} \end{cases}$$

Replications



Reminder: Central Limit Theorem

Let $X_1, X_2, \dots, X_n$ be a set of n independent and identically distributed random variables with mean μ and a finite variance σ^2. Let the sample average be defined as

$$S_n := \frac{X_1 + \cdots + X_n}{n}$$

As $n \to \infty$

$$S_n \to \mathcal{N}(\mu, \sigma^2/n)$$

Regardless of distribution of X_i!

Replications

- The outputs from a simulation model are random variables
- Running the simulator provides realizations of these r.v.
- Replications allow us to:
  - Obtain independent observations.
    This allows us to apply classical statistical methods to analyze the outputs:
    e.g. central limit theorem, confidence intervals
  - Estimate system performance measures (e.g. empirical cdf) that enable us to understand the "typical" behavior of the system
  - Have an idea of the underlying (unknown and often) complex distribution of the output variable
- How to obtain different replications with a simulation software?
  - For a given replication the sequence of random numbers are generated starting with an initial number, called the seed
  - Each time you launch the simulation with the same seed, you obtain identical results
  - Saving/knowing the seed, allows you to reproduce your results
  - Launching the simulation with different seeds, yields different realizations

Statistical issues

- The statistical analysis of the results of a simulation can be difficult
- Impact of initial conditions (warm-up period, before reaching a steady state)
- Selection of starting conditions
- Correlation between successive observations/samples (e.g., waiting times of passengers in a system)
- Number and length of required replications:
  - Many replications vs. single long replication (which is then partitioned)
- Confidence intervals
- Statistical tests

Simulation project

The bigger picture:

- Problem definition
- Model formulation
- Data collection
- **Model development**
- Verification
- Validation
- Experiments (e.g. what-if analysis)
- Results: analysis and presentation
- Implementation

Outline

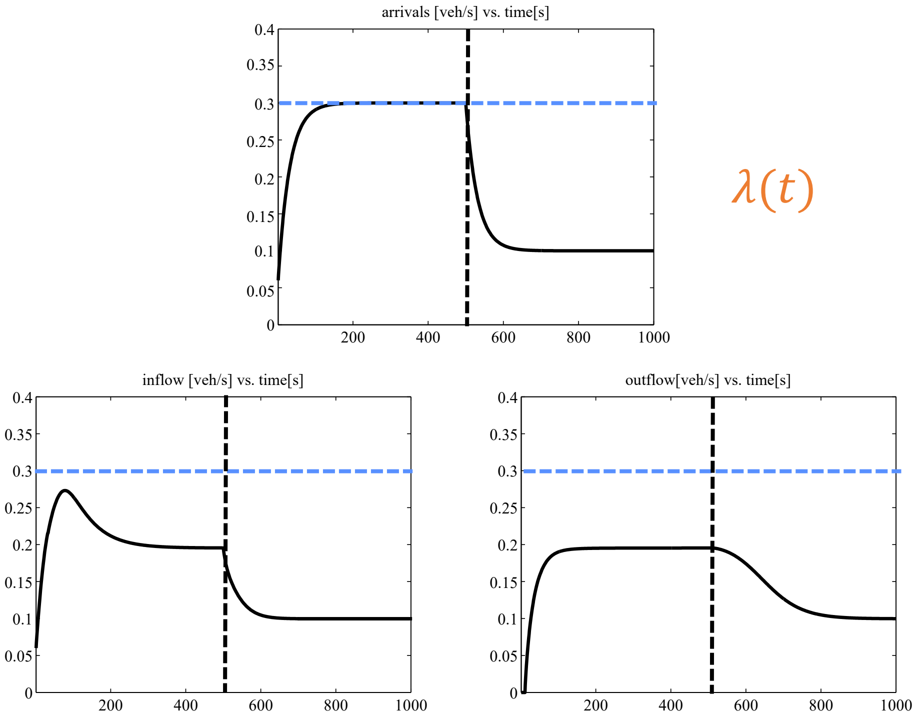
1. Introduction to simulation
2. Discrete event simulation
3. **Transient analysis**
4. Mixed continuous and discrete event simulation

Transient analysis for urban traffic

Dynamic analysis of one link over 1000 s

- Initially empty link, arrival rate that is 0.3 veh/s for the first 500 s and then jumps down to 0.1 veh/s, where it stays for the remaining 500 s
- The downstream flow capacity (service rate) of the link is 0.2 veh/s
  - Thus the first half of the demand exceeds the link's bottleneck capacity, whereas the second half can be served
- Dynamic analysis of queue build-up and dissipation

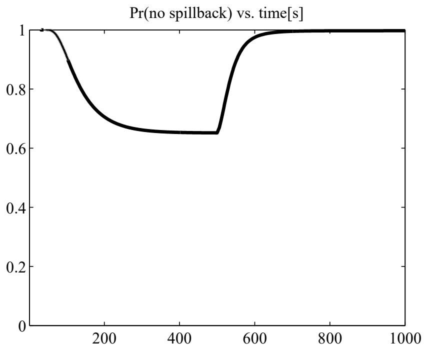
Transient analysis for urban traffic

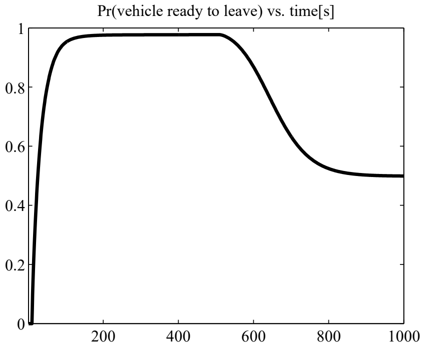


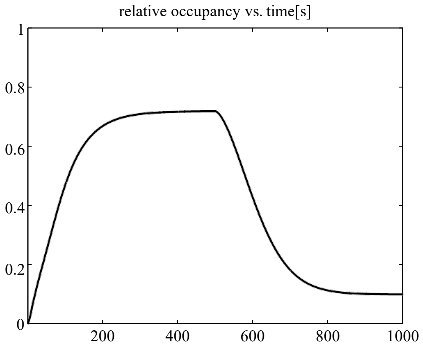
$\lambda(t)$

- (1) Arrivals, (2) Inflow, (3) Outflow

Transient analysis for urban traffic







1. Probability that there is no spillback
2. Probability that a vehicle is ready to leave the link
3. Relative occupancy (expected number of vehicles divided by maximum number of vehicles)

Transient analysis

- Relax the assumption of stationarity
- Analyze the transient (dynamic) behavior of a queue
- Transition rate (linear) differential equations for a birth-death process:

$$\begin{cases} \dfrac{dP_n(t)}{dt} = -(\lambda_n + \mu_n)P_n(t) + \mu_{n+1}P_{n+1}(t) + \lambda_{n-1}P_{n-1}(t), \forall n \geq 1 \\ \dfrac{dP_0(t)}{dt} = -\lambda_0 P_0(t) + \mu_1 P_1(t) \end{cases}$$

- Can be written as:

$$\frac{dP(t)}{dt} = P(t)Q$$

- This linear system of differential equations has general solution:

$$P(t) = P(0)e^{Qt}$$

- Numerical methods used to evaluate $P(t)$

Transient analysis

- M/M/1/K queue

$$\begin{cases} P_n(t) = s_n + \rho^{\frac{n}{2}} \sum_{j=1}^{K} C_j \left(\sin \frac{jn\pi}{K+1} - \sqrt{\rho} \sin \frac{j(n+1)\pi}{K+1} \right) e^{\tau_j t} \\ \tau_j = \lambda + \mu - 2\sqrt{\lambda\mu} \cos \frac{j\pi}{K+1} \end{cases}$$

where s is the stationary dbn, and the coefficients $\{C_j\}$ are chosen to fit the initial values of the transient distribution.

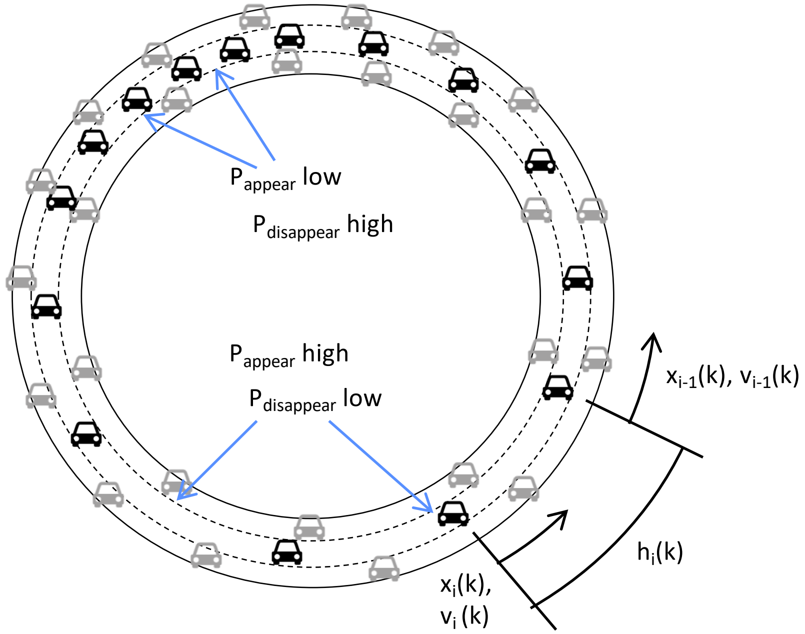
- There are few closed-form expressions for transient distributions

Outline

1. Introduction to simulation
2. Discrete event simulation
3. Transient analysis
4. **Mixed continuous and discrete event simulation**

(Simple) multi-lane modeling

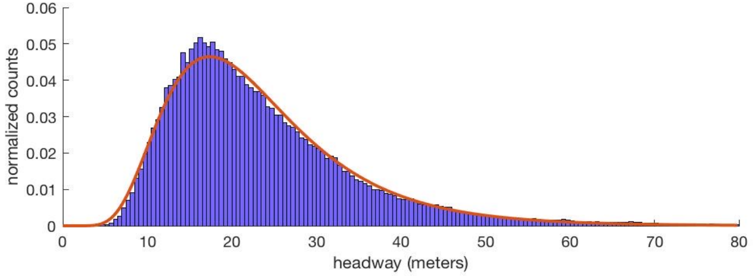
Stochastic single-lane model

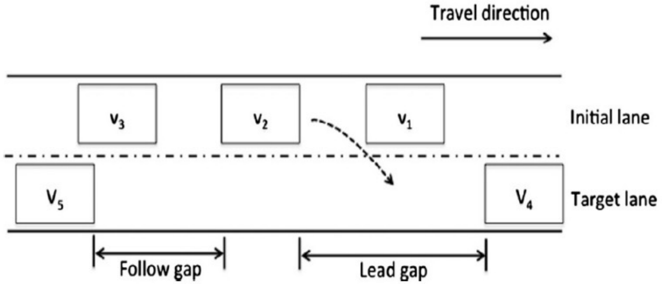


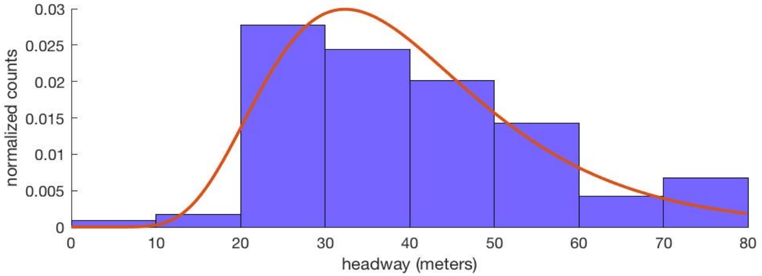
Wu, Cathy, Eugene Vinitsky, Aboudy Kreidieh, and Alexandre Bayen. "Multi-lane reduction: A stochastic single-lane model for lane changing." IEEE ITSC, 2017.

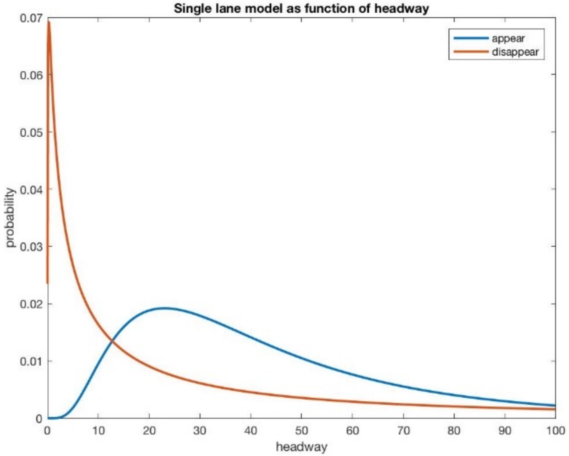
Model Calibration - Probabilities

Extract headways when lane changes occur
Fit to log-normal distribution







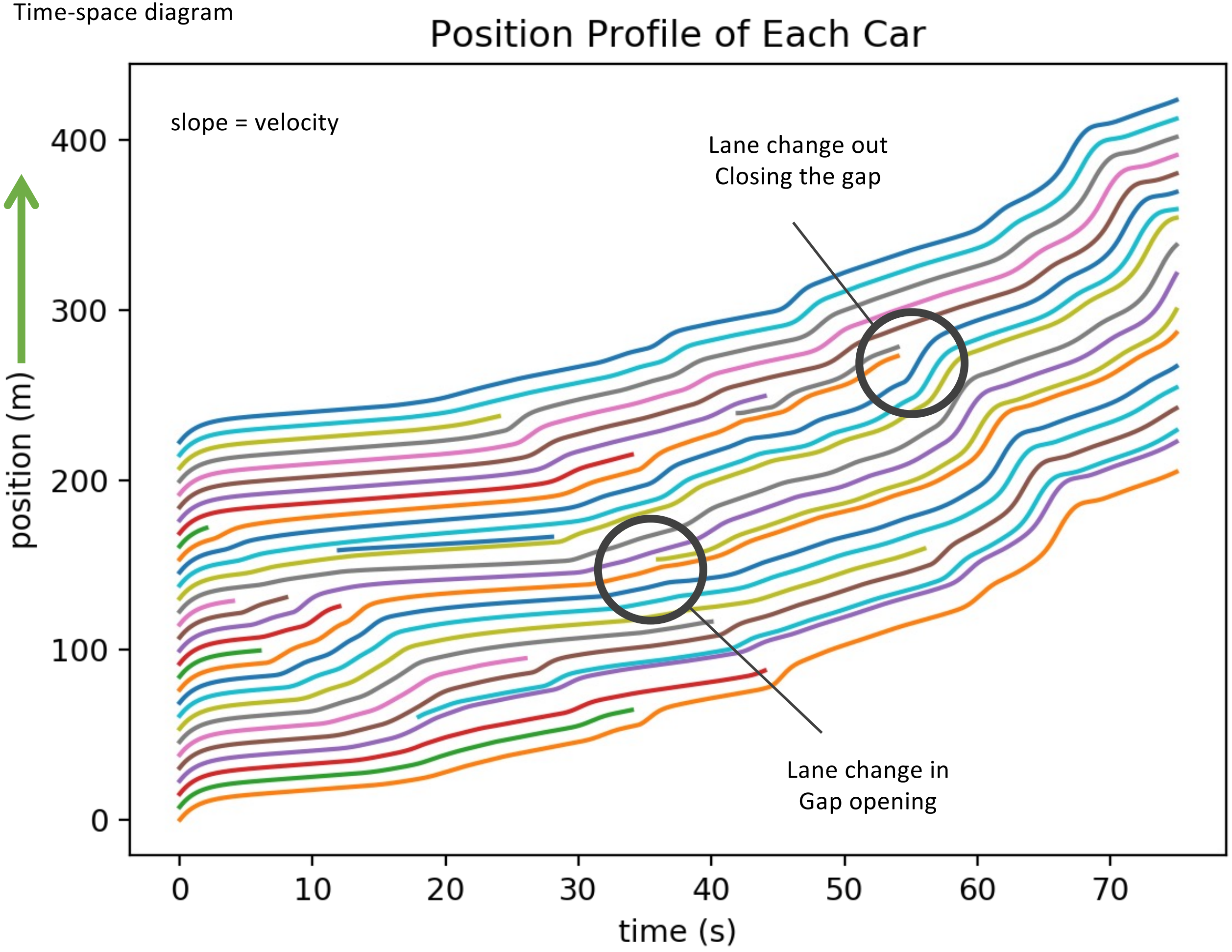


Fit of log-normal distribution for the total distribution of headways and the conditional distribution of headways when a vehicle lane changes into a lane, respectively, computed for 7:50 am on the US 101.

Wu, Cathy, Eugene Vinitsky, Aboudy Kreidieh, and Alexandre Bayen. "Multi-lane reduction: A stochastic single-lane model for lane changing." IEEE ITSC, 2017.

Slide credit: Eugene Vinitsky

Time-space diagram



Wu, Cathy, Eugene Vinitsky, Aboudy Kreidieh, and Alexandre Bayen. "Multi-lane reduction: A stochastic single-lane model for lane changing." IEEE ITSC, 2017.

References

1. Larson, Richard C. and Amedeo R. Odoni. **Urban Operations Research**. Prentice-Hall (1981). Chapter 7: Simulations.
2. Slides adapted from Carolina Osorio