MTAT.03.231
Business Process Management

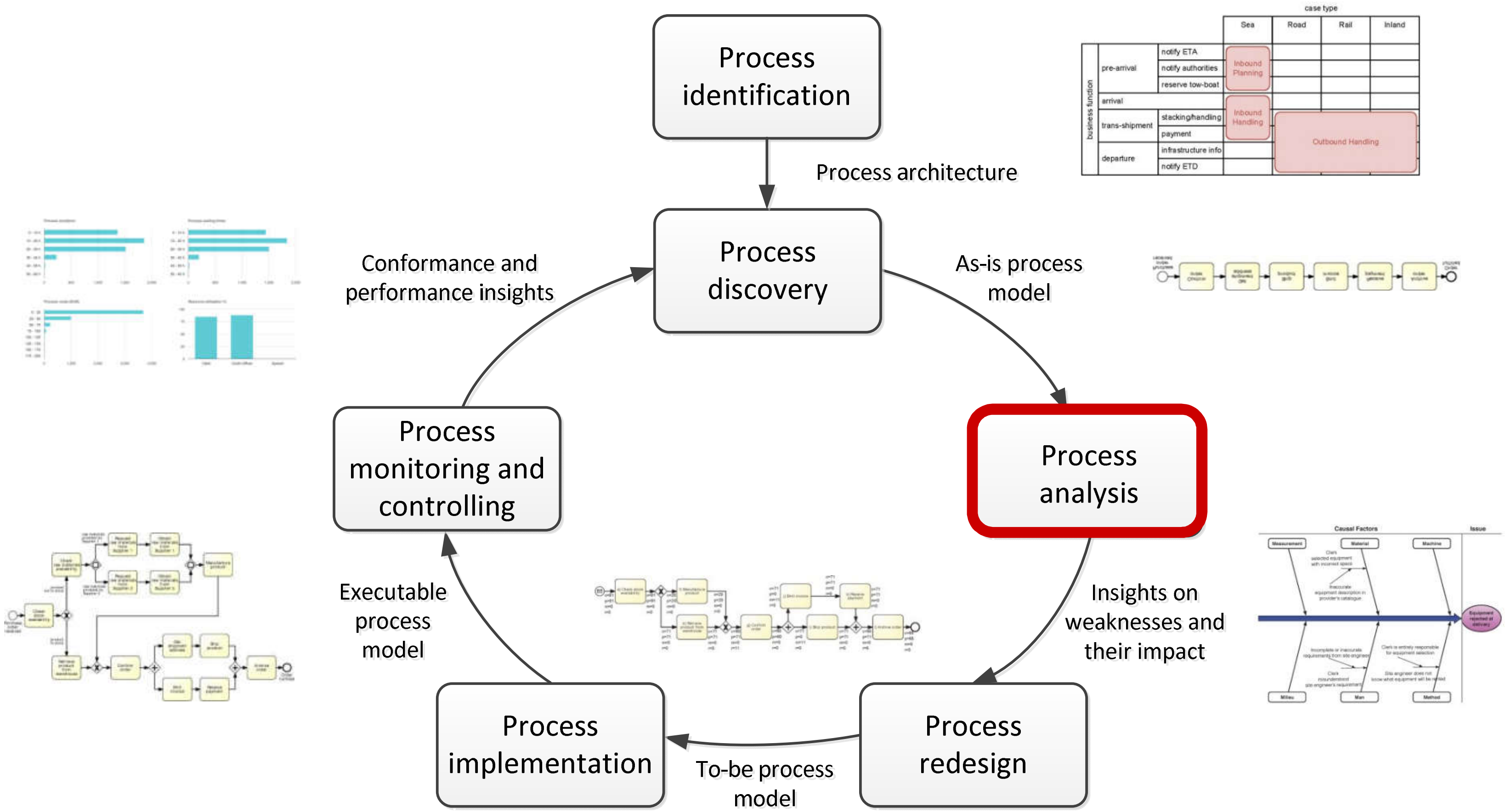
Lecture 7 – Quantitative Process Analysis II

Marlon Dumas

marlon.dumas ät ut . ee

UNIVERSITY OF TARTU
INSTITUTE OF COMPUTER SCIENCE

Process Analysis



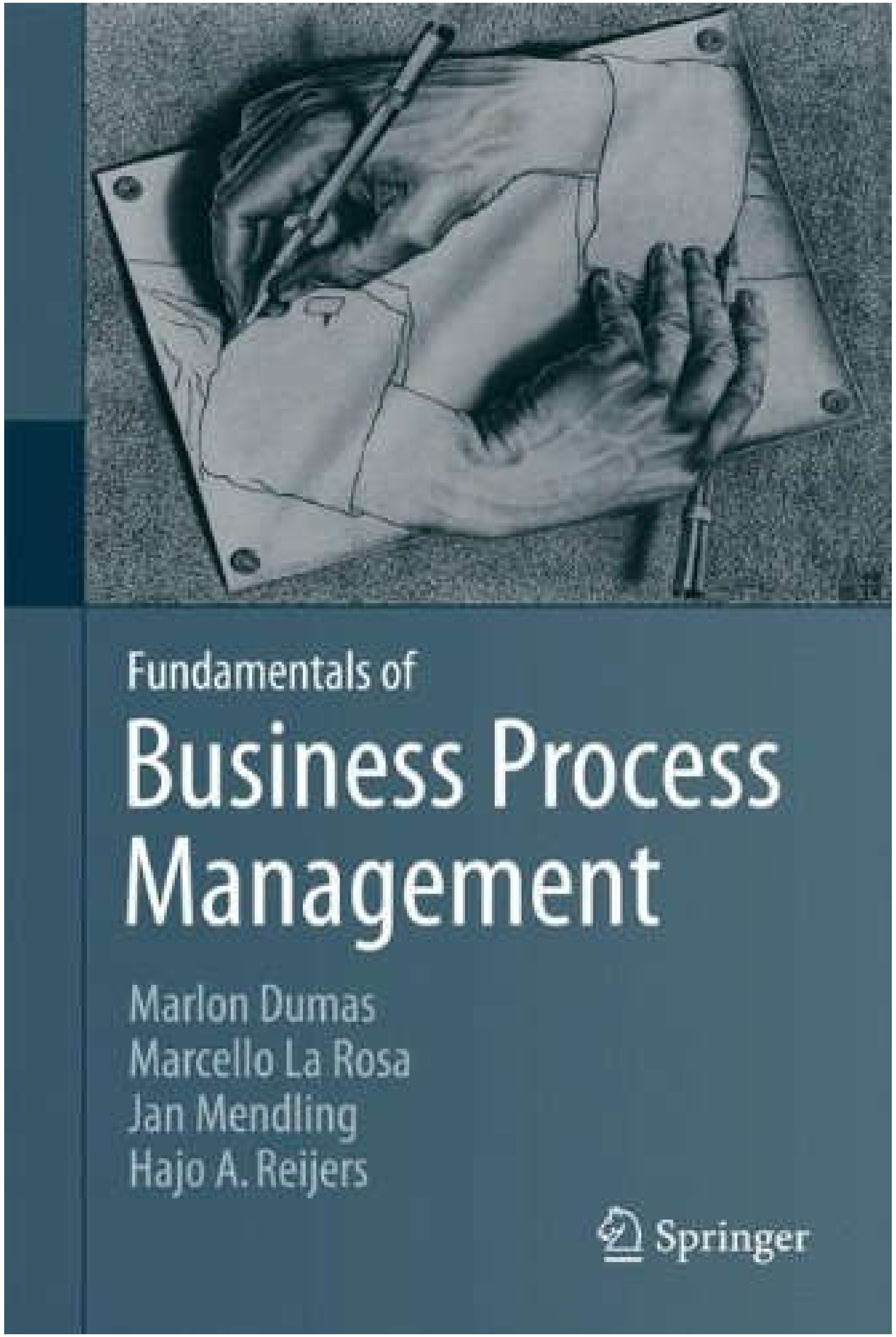
Process Analysis Techniques

Qualitative analysis

- Value-Added & Waste Analysis
- Root-Cause Analysis
- Pareto Analysis
- Issue Register

Quantitative Analysis

- Flow analysis
- Queuing analysis
- Simulation



1. Introduction
2. Process Identification
3. Essential Process Modeling
4. Advanced Process Modeling
5. Process Discovery
6. Qualitative Process Analysis
7. **Quantitative Process Analysis**
8. Process Redesign
9. Process Automation
10. Process Intelligence

Why flow analysis is not enough?

Flow analysis does not consider waiting times due to resource contention

Queuing analysis and simulation address these limitations and have a broader applicability

Queuing Analysis

- Capacity problems are common and a key driver of process redesign
  - Need to balance the cost of increased capacity against the gains of increased productivity and service
- Queuing and waiting time analysis is particularly important in service systems
  - Large costs of waiting and/or lost sales due to waiting

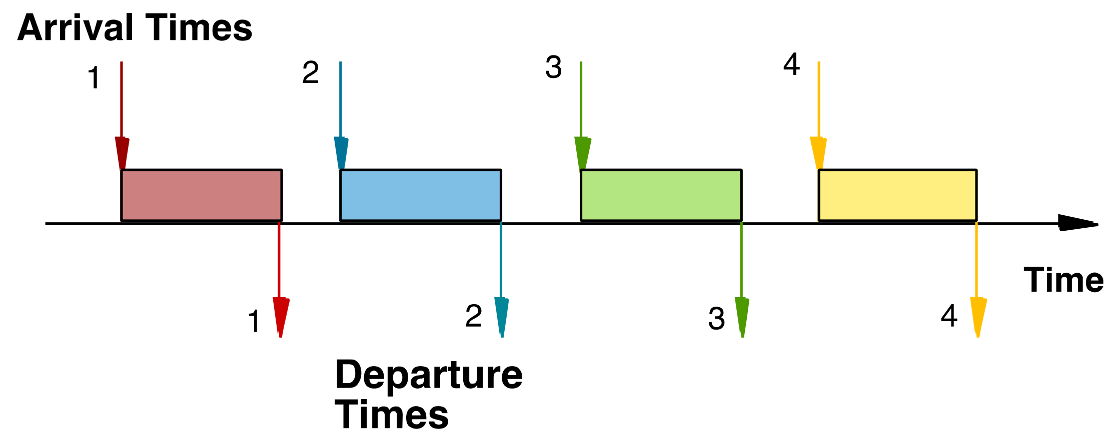
Prototype Example – ER at a Hospital

- Patients arrive by ambulance or by their own accord
- One doctor is always on duty
- More patients seeks help $\Rightarrow$ longer waiting times

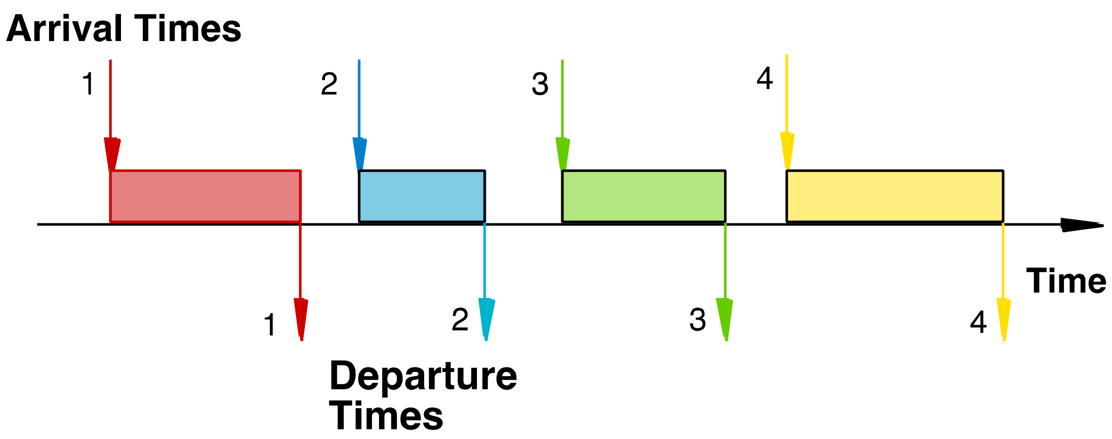
➢ ***Question**: Should another MD position be instated?*

Delay is Caused by Job Interference

If arrivals are regular or sufficiently spaced apart, no queuing delay occurs



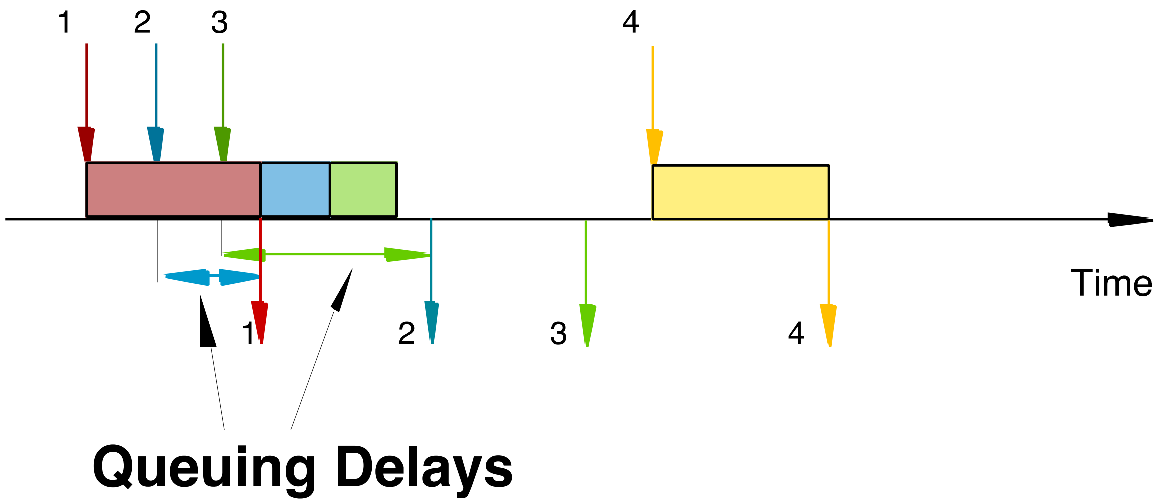
Deterministic traffic



Variable but spaced apart traffic

Burstiness Causes Interference

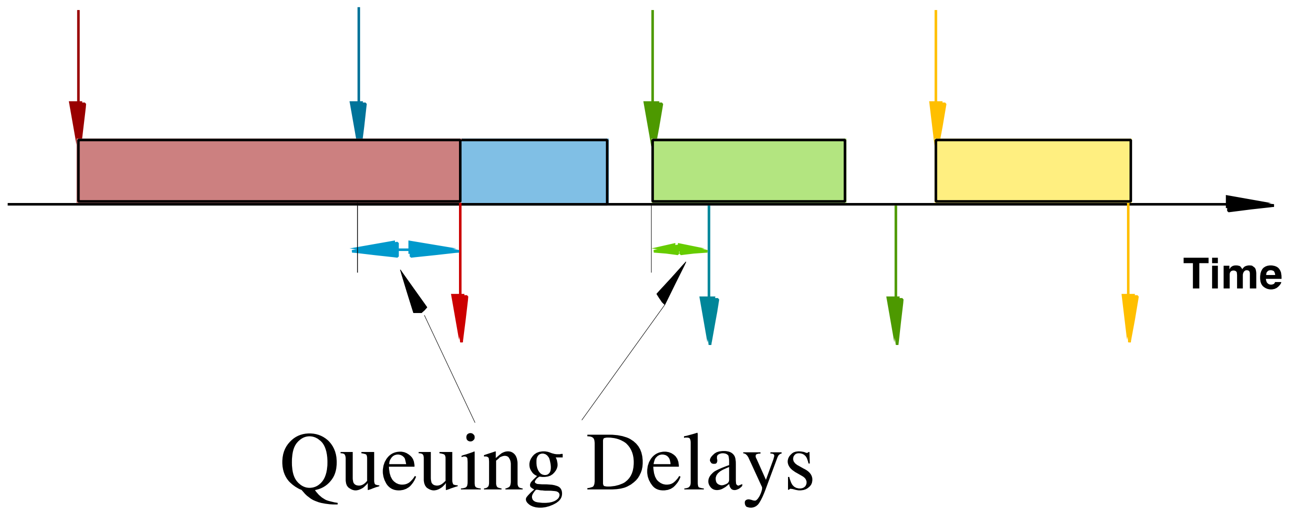
- Queuing results from variability in processing times and/or interarrival intervals



Bursty Traffic

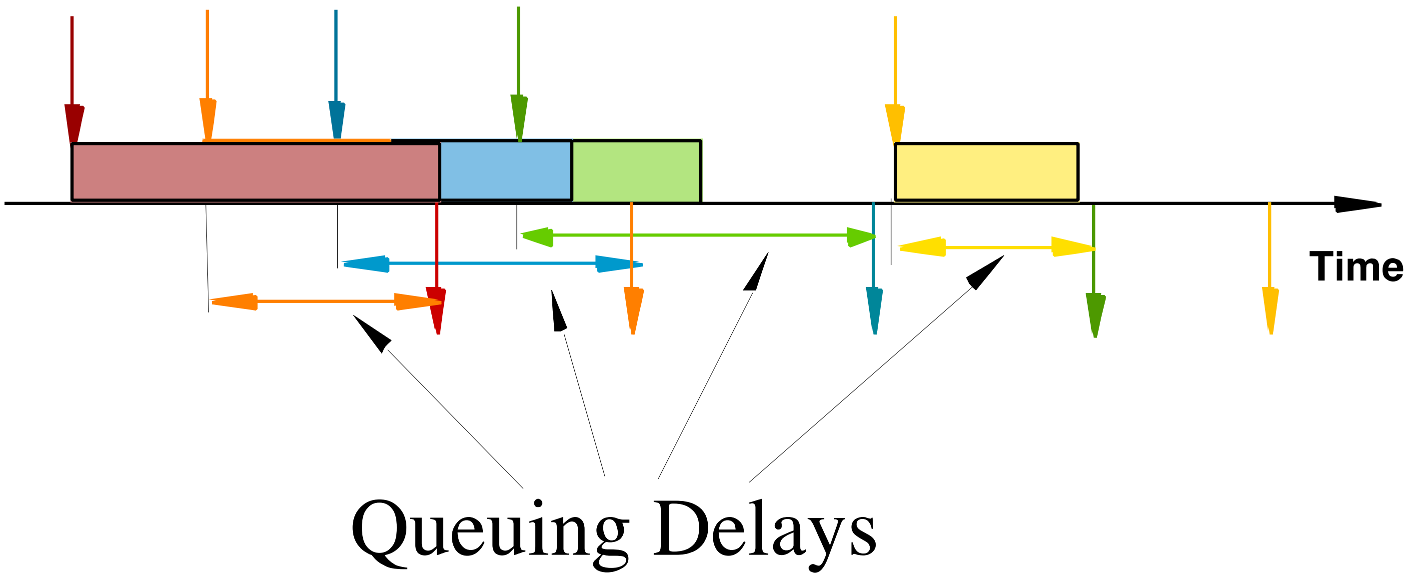
Job Size Variation Causes Interference

- Deterministic arrivals, variable job sizes



High Utilization Exacerbates Interference

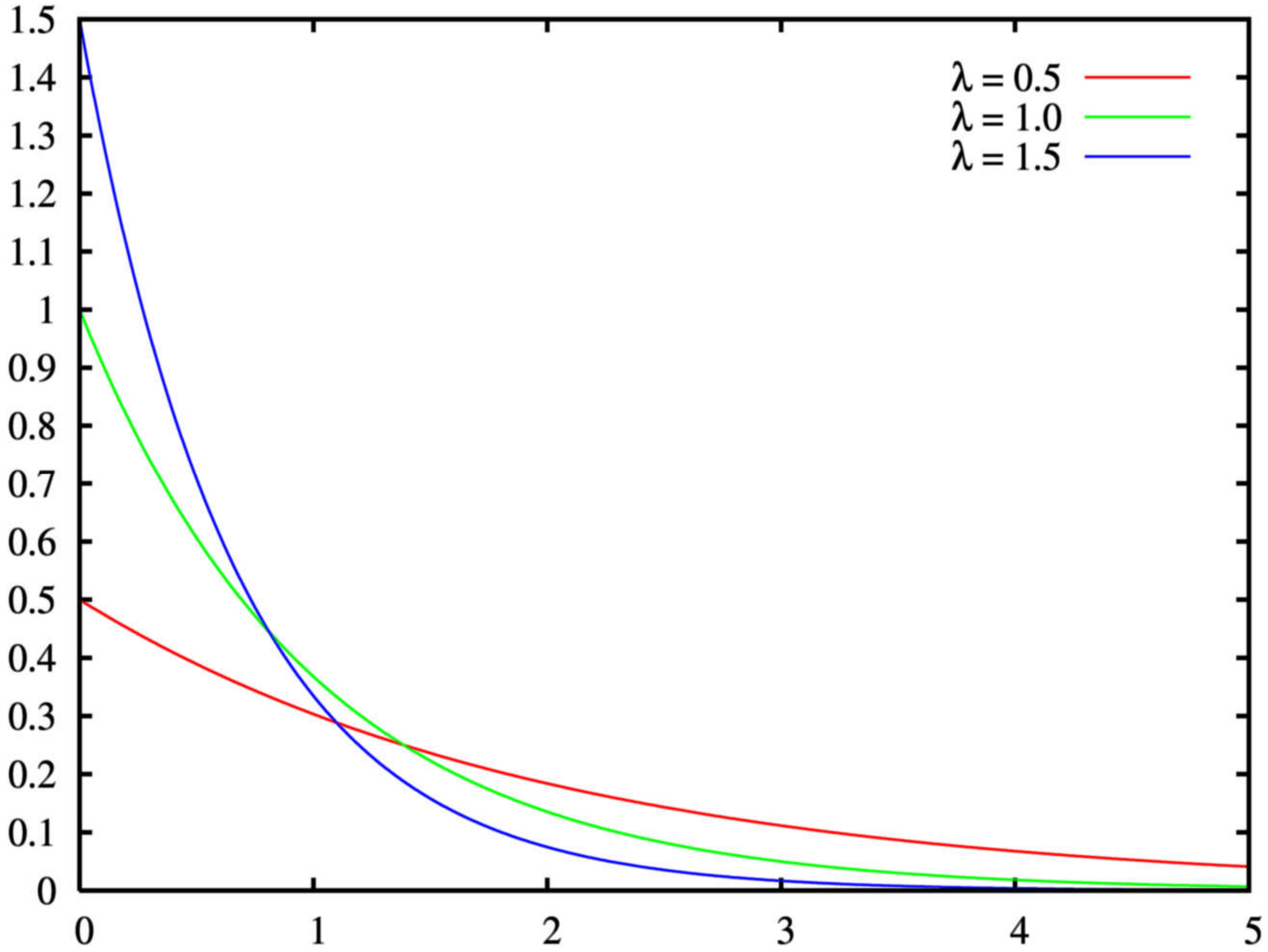
- The queuing probability increases as the load increases
- Utilization close to 100% is unsustainable → too long queuing times



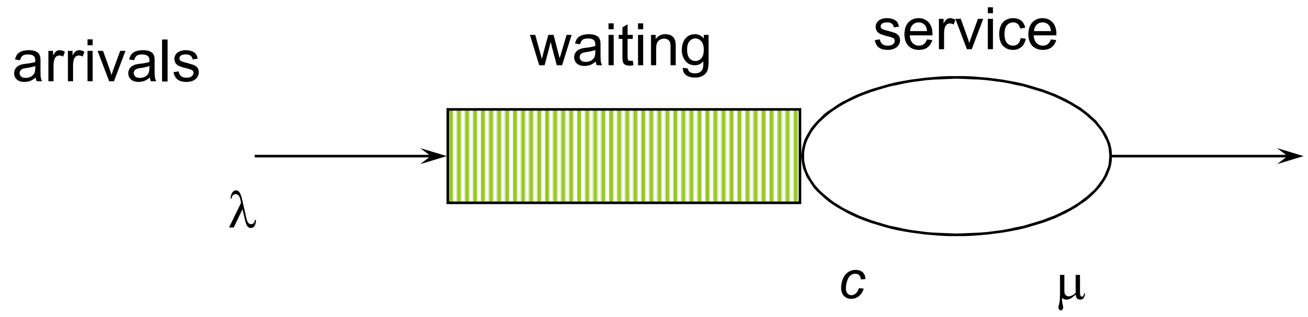
The Poisson Process

- Common arrival assumption in many queuing and simulation models
- The times between arrivals are independent, identically distributed and exponential
  - P (arrival < t) = $1 - e^{-\lambda t}$
- Key property: The fact that a certain event has not happened tells us nothing about how long it will take before it happens
  - e.g., $P(X > 40 \mid X >= 30) = P(X > 10)$

Negative Exponential Distribution



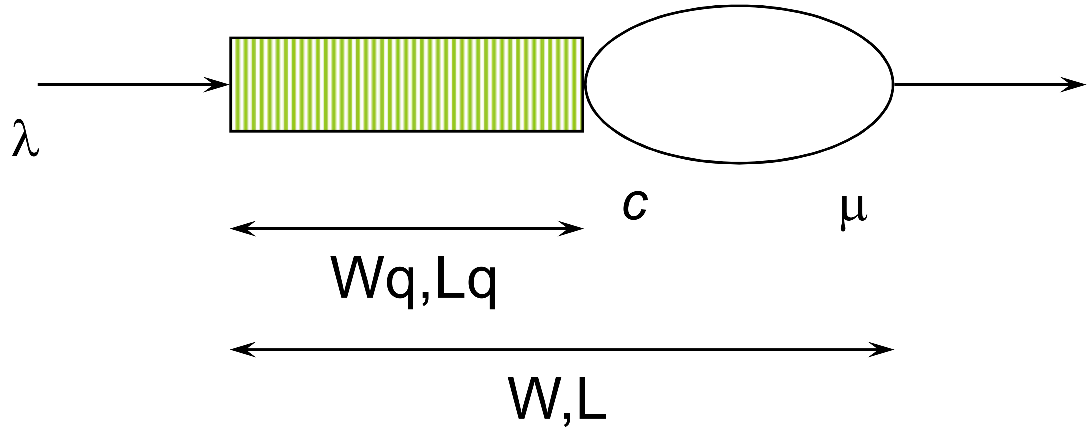
Queuing theory: basic concepts



Basic characteristics:

- λ (mean arrival rate) = average number of arrivals per time unit
- μ (mean service rate) = average number of jobs that can be handled by one server per time unit:
- c = number of servers

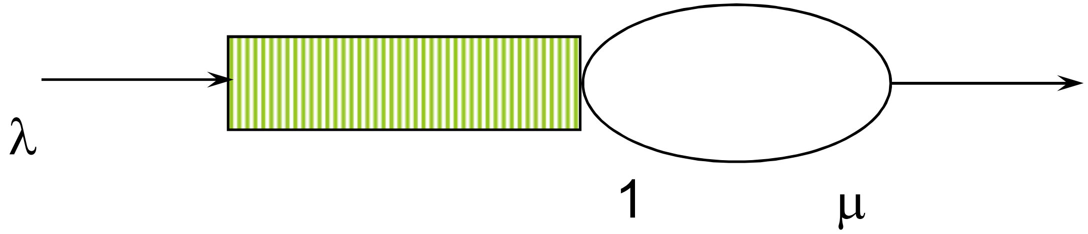
Queuing theory concepts (cont.)



Given λ , μ and c, we can calculate :

- $\rho =$ **resource utilization**
- **Wq = average time a job spends in queue (i.e. waiting time)**
- **W = average time in the "system" (i.e. *cycle time*)**
- **Lq = average number of jobs in queue (i.e. length of queue)**
- **L = average number of jobs in system (i.e. *Work-in-Progress*)**

M/M/1 queue



Assumptions:

- time between arrivals and processing time follow a negative exponential distribution
- 1 server (c = 1)
- FIFO

$$\rho = \frac{\text{Capacity Demand}}{\text{Available Capacity}} = \frac{\lambda}{\mu}$$

$$L = \rho/(1-\rho)$$

$$W = L/\lambda = 1/(\mu - \lambda)$$

$$L_q = \rho^2/(1-\rho) = L - \rho$$

$$W_q = L_q/\lambda = \lambda/(\mu(\mu-\lambda))$$

M/M/c queue

- Now there are c servers in parallel, so the expected capacity per time unit is then $c^*\mu$

$$\rho = \frac{\text{Capacity Demand}}{\text{Available Capacity}} = \frac{\lambda}{c * \mu}$$

Little's Formula $\Rightarrow$ $\mathbf{W_q = L_q / \lambda}$

$\mathbf{W = W_q + (1/\mu)}$

Little's Formula $\Rightarrow$ $\mathbf{L = \lambda W}$

Tool Support

- For M/M/c systems, the exact computation of L_q is rather complex…

$$L_q = \sum_{n=c}^{\infty}(n-c)P_n = \ldots = \frac{(\lambda/\mu)^c \rho}{c!(1-\rho)^2} P_0$$

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^c}{c!} \cdot \frac{1}{1-(\lambda/(c\mu))} \right)^{-1}$$

- Consider using a tool, e.g.
  - http://www.supositorio.com/rcalc/rcalclite.htm (very simple)
  - http://queueingtoolpak.org/ (more sophisticated, Excel add-on)

Example – ER at County Hospital

- Situation
  - Patients arrive according to a Poisson process with intensity λ ($\Leftrightarrow$ the time between arrivals is exp(λ) distributed.
  - The service time (the doctor's examination and treatment time of a patient) follows an exponential distribution with mean $1/\mu$ (=exp(μ) distributed)

  $\Rightarrow$ *The ER can be modeled as an M/M/c system where c = the number of doctors*

- Data gathering

  $\Rightarrow$ λ = 2 patients per hour

  $\Rightarrow$ μ = 3 patients per hour
- Question
  - Should the capacity be increased from 1 to 2 doctors?

Queuing Analysis – Hospital Scenario

- Interpretation
  - To be in the queue = to be in the waiting room
  - To be in the system = to be in the ER (waiting or under treatment)

Characteristic	One doctor (c=1)	Two Doctors (c=2)
ρ	2/3	1/3
L_q	4/3 patients	1/12 patients
L	2 patients	3/4 patients
W_q	2/3 h = 40 minutes	1/24 h = 2.5 minutes
W	1 h	3/8 h = 22.5 minutes

- Is it warranted to hire a second doctor ?

Your turn

- Textbook, Chapter 7, exercise 7.12

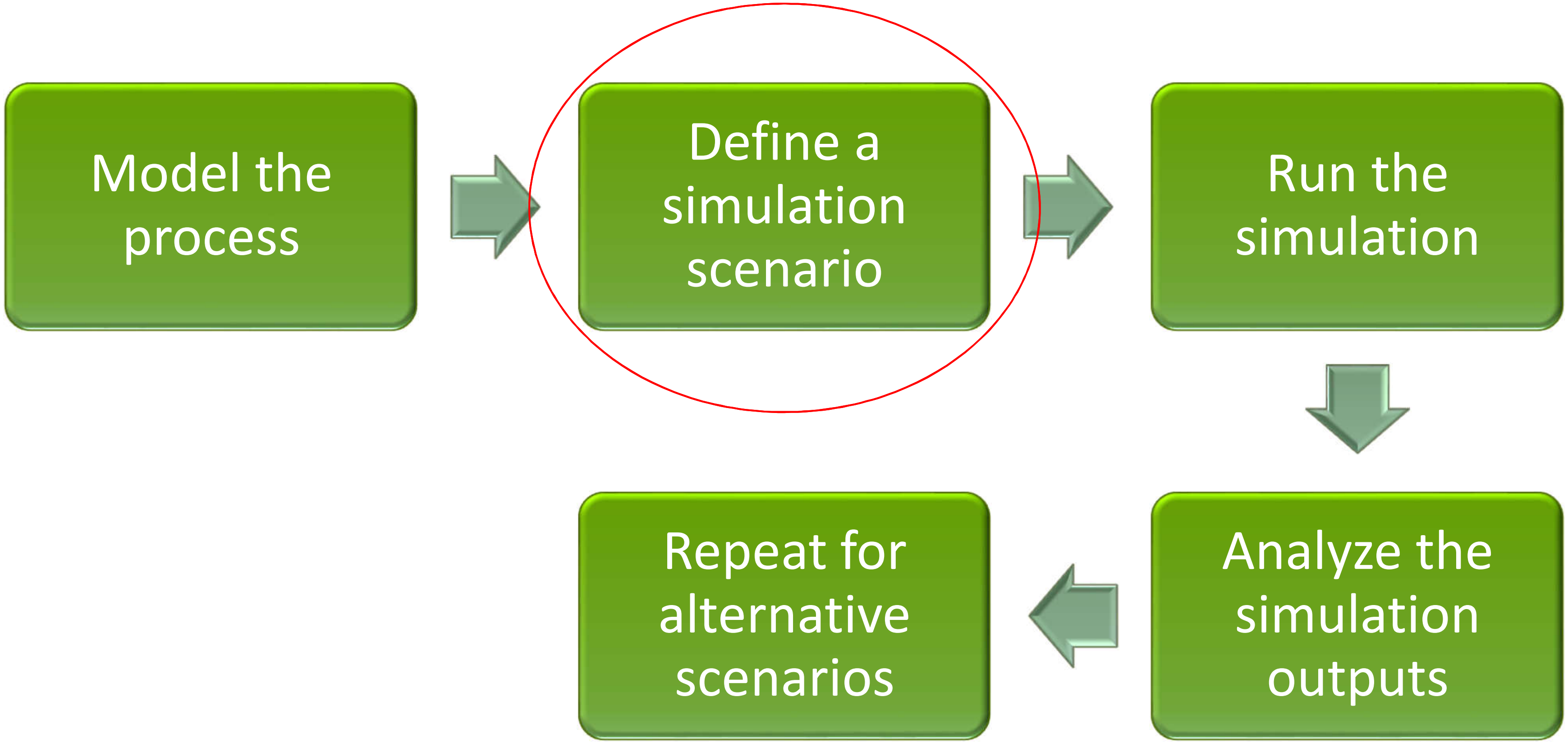
We consider a Level-2 IT service desk with two staff members. Each staff member can handle one service request in 4 working hours on average. Service times are exponentially distributed. Requests arrive at a mean rate of one request every 3 hours according to a Poisson process. What is the average time between the moment a service request arrives at this desk and the moment it is fulfilled?

…Now consider the scenario where the number of requests becomes one per hour. How many level-2 staff are required to be able to start serving a request on average within two working hours of it being received?

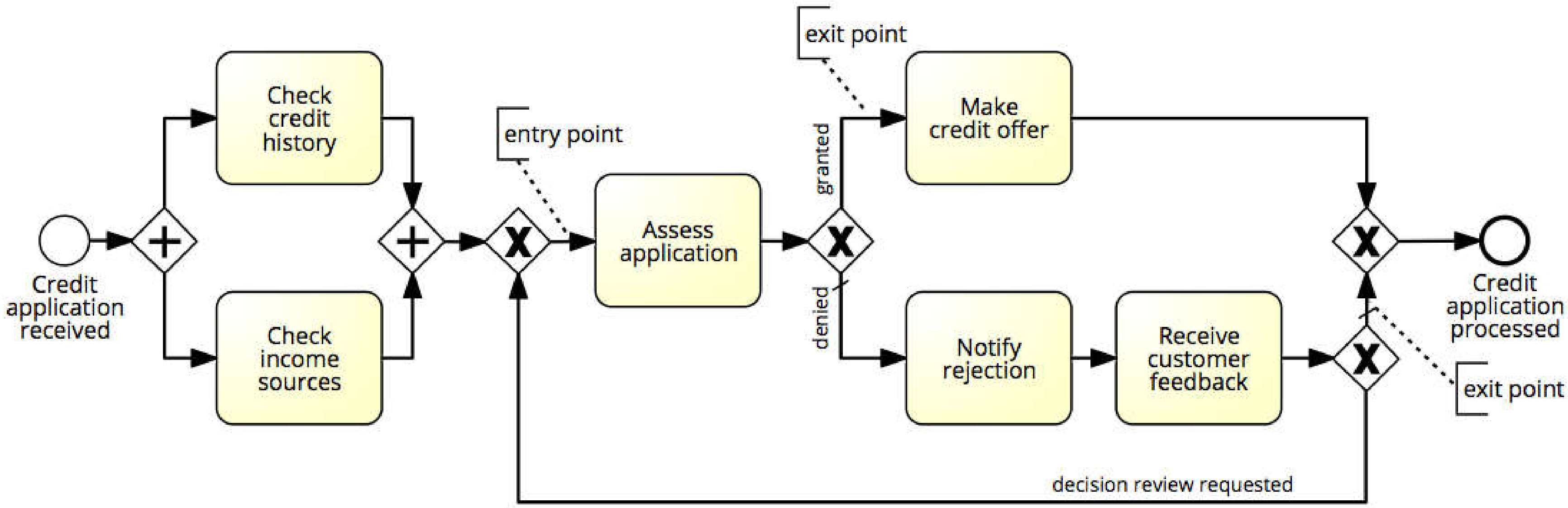
Process Simulation

- Versatile quantitative analysis method for
  - As-is analysis
  - What-if analysis
- In a nutshell:
  - Run a large number of process instances
  - Gather performance data (cost, time, resource usage)
  - Calculate statistics from the collected data

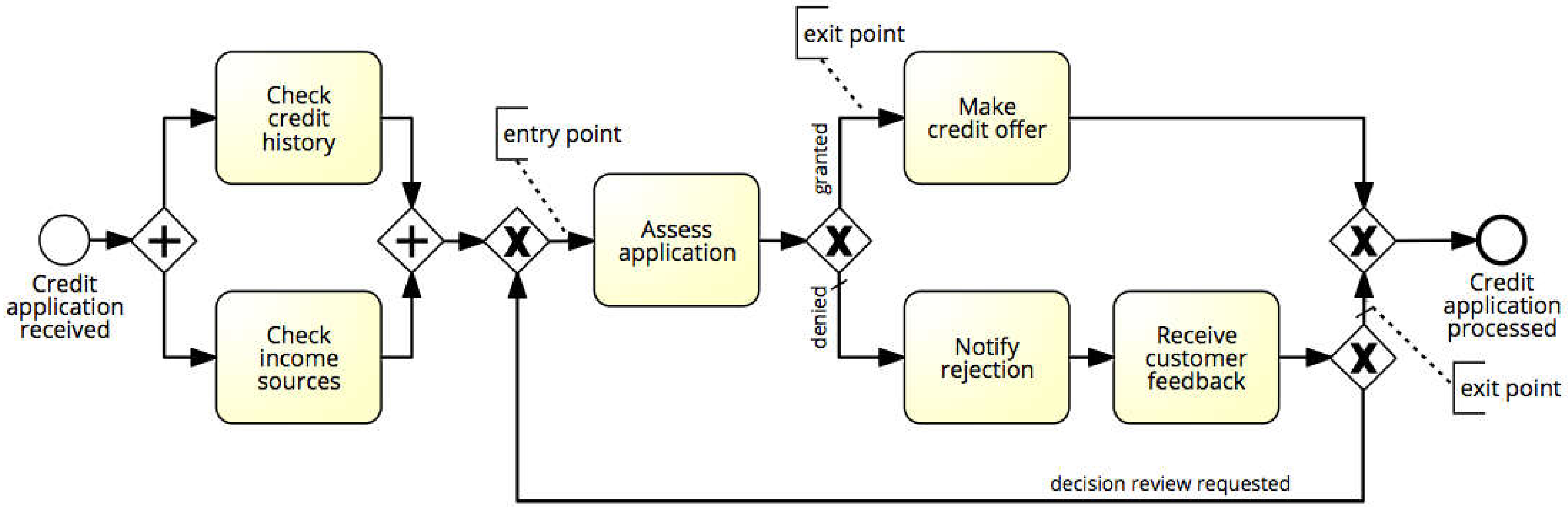
Process Simulation



Example


exit point
entry point
Check credit history
Make credit offer
Assess application
granted
denied
Credit application received
Check income sources
Notify rejection
Receive customer feedback
Credit application processed
exit point
decision review requested

Example

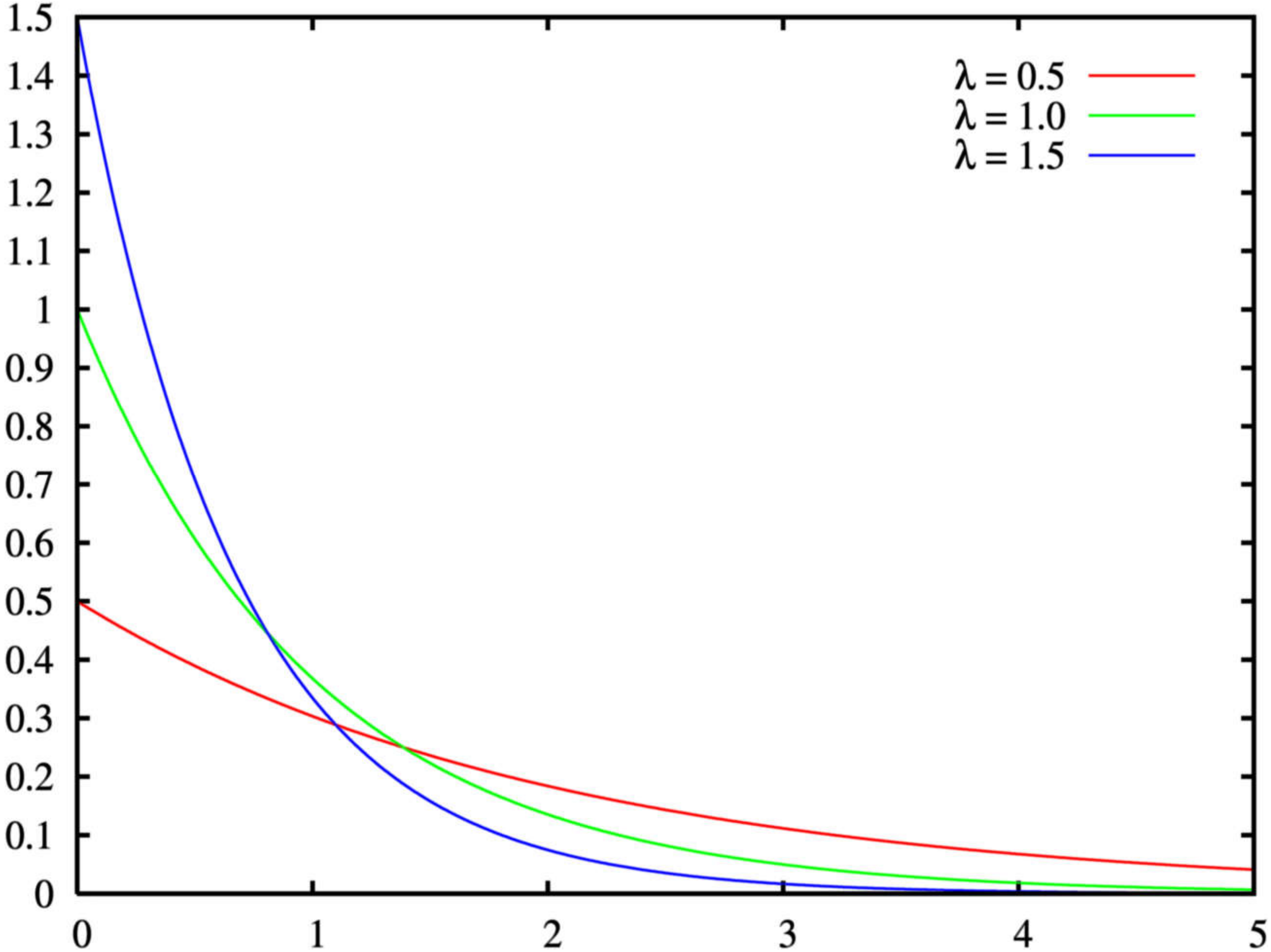


Check credit history
Check income sources
Credit application received
entry point
Assess application
exit point
granted
denied
Make credit offer
Notify rejection
Receive customer feedback
exit point
Credit application processed
decision review requested

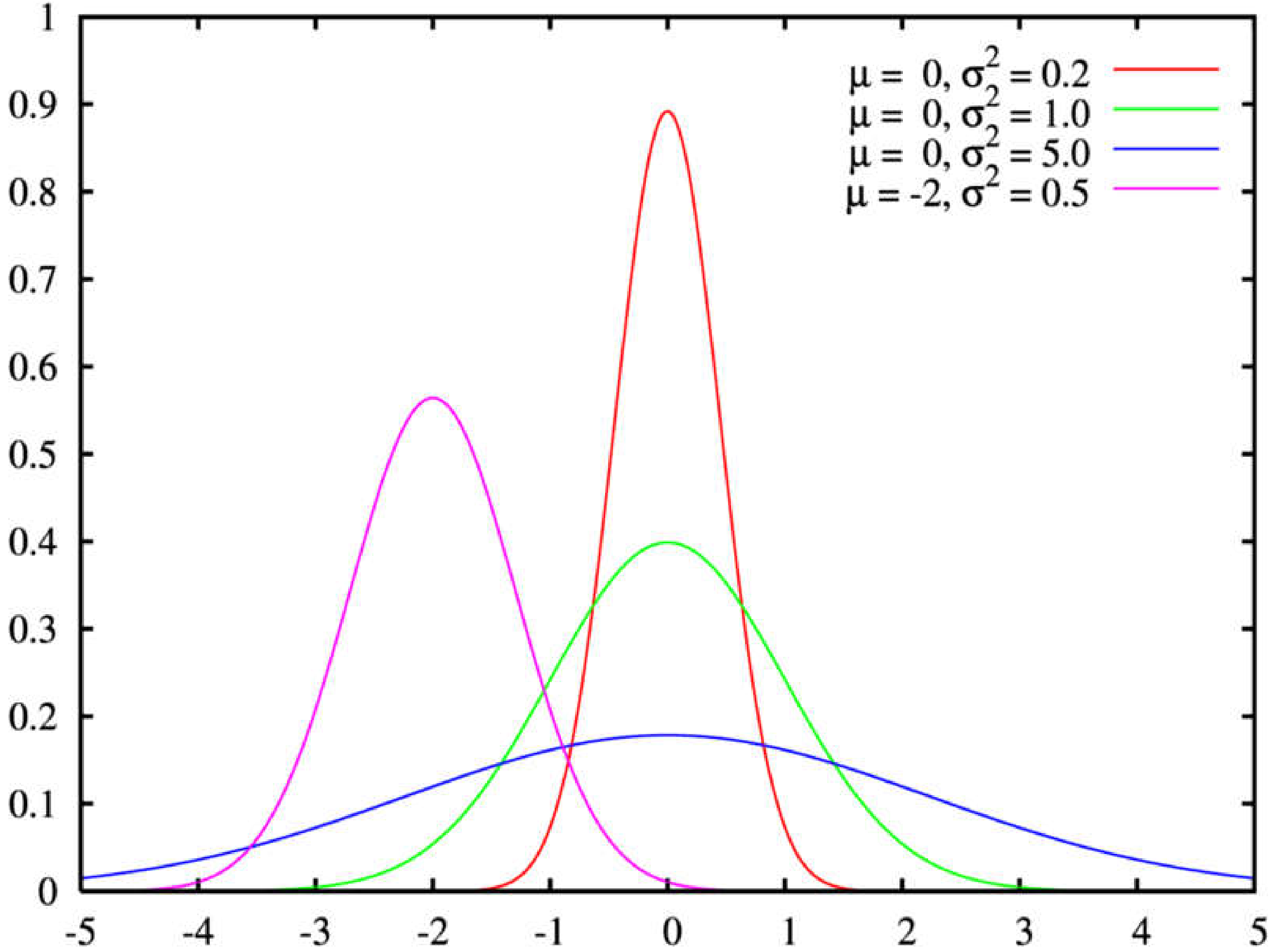
Elements of a simulation scenario

1. Processing times of activities
   - Fixed value
   - Probability distribution

Exponential Distribution



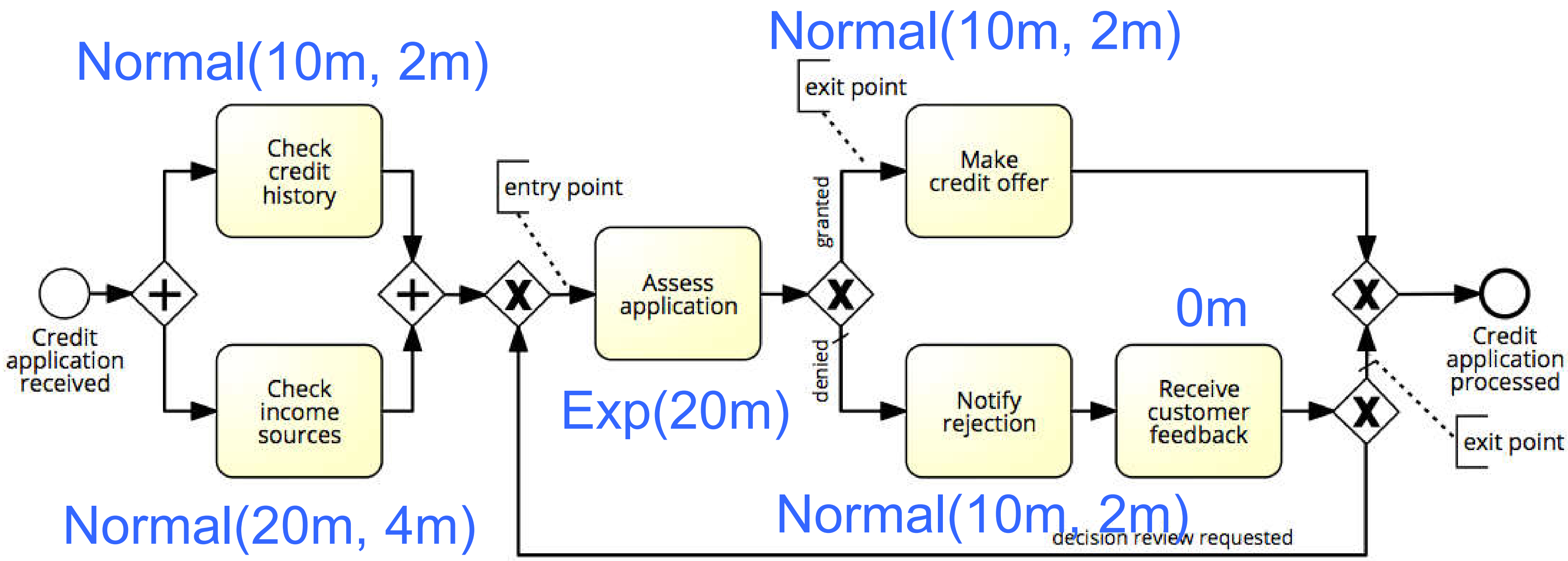
Normal Distribution



Choice of probability distribution

- Fixed
  - Rare, can be used to approximate case where the activity processing time varies very little
  - Example: a task performed by a software application
- Normal
  - Repetitive activities
  - Example: "Check completeness of an application"
- Exponential
  - Complex activities that may involve analysis or decisions
  - Example: "Assess an application"

Simulation Example

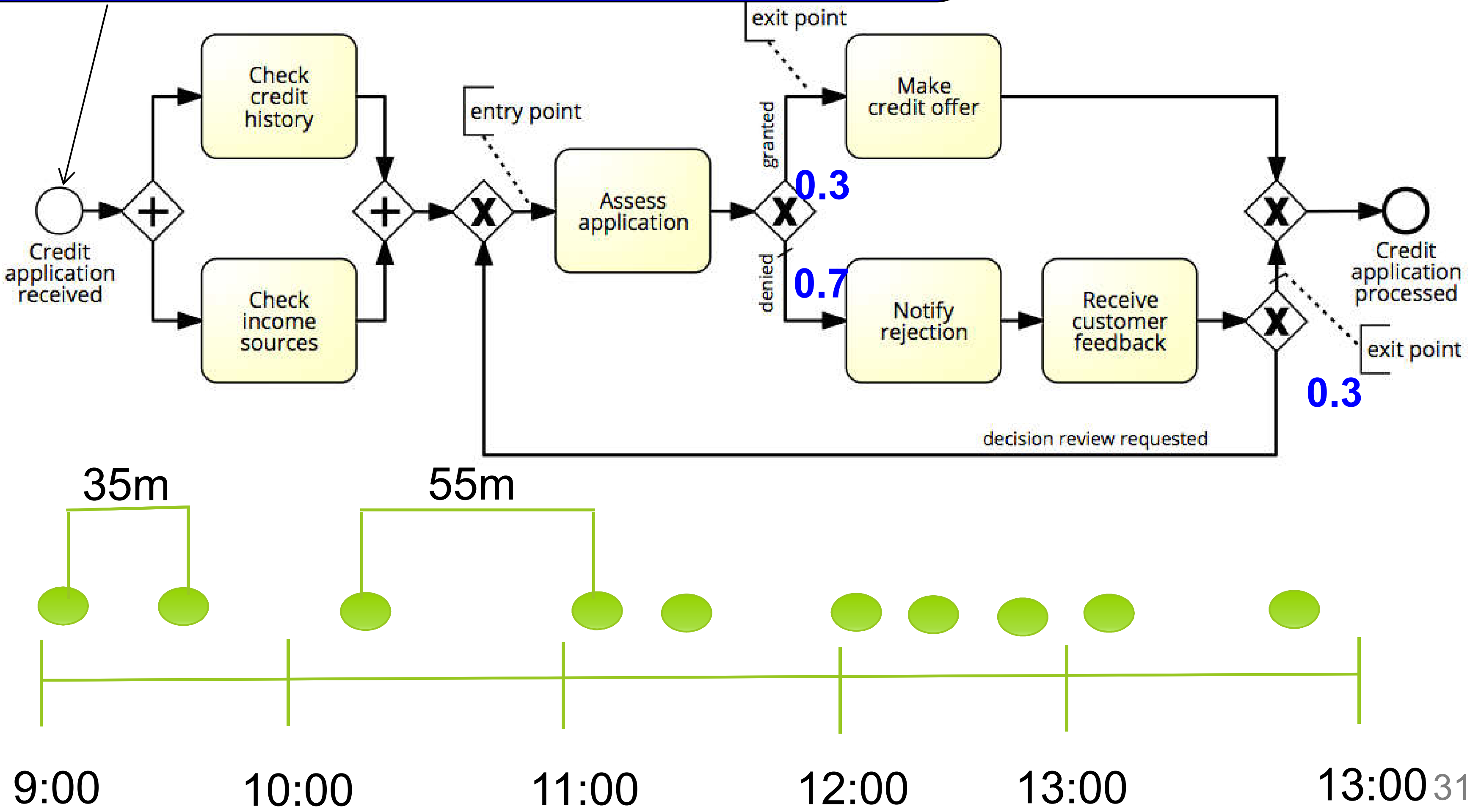


Elements of a simulation model

1. Processing times of activities
   - Fixed value
   - Probability distribution
2. Conditional branching probabilities
3. Arrival rate of process instances and probability distribution
   - Typically exponential distribution with a given mean inter-arrival time
   - Arrival calendar, e.g. Monday-Friday, 9am-5pm, or 24/7

Branching probability and arrival rate

Arrival rate = 2 applications per hour
Inter-arrival time = 0.5 hour
Negative exponential distribution
From Monday-Friday, 9am-5pm



Elements of a simulation model

1. Processing times of activities
   - Fixed value
   - Probability distribution
2. Conditional branching probabilities
3. Arrival rate of process instances
   - Typically exponential distribution with a given mean inter-arrival time
   - Arrival calendar, e.g. Monday-Friday, 9am-5pm, or 24/7
4. Resource pools

Resource pools

- Name
- Size of the resource pool
- Cost per time unit of a resource in the pool
- Availability of the pool (working calendar)
- Examples

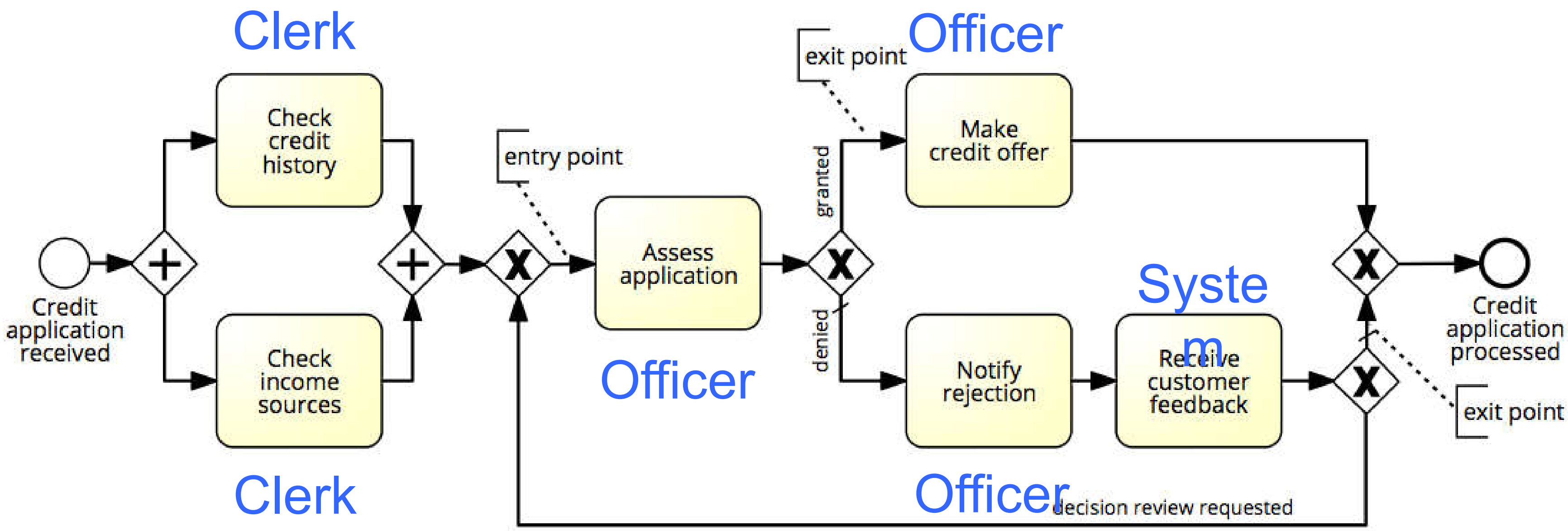
Clerk	Credit Officer
€ 25 per hour	€ 25 per hour
Monday-Friday, 9am-5pm	Monday-Friday, 9am-5pm

- In some tools, it is possible to define cost and calendar per resource, rather than for entire resource pool

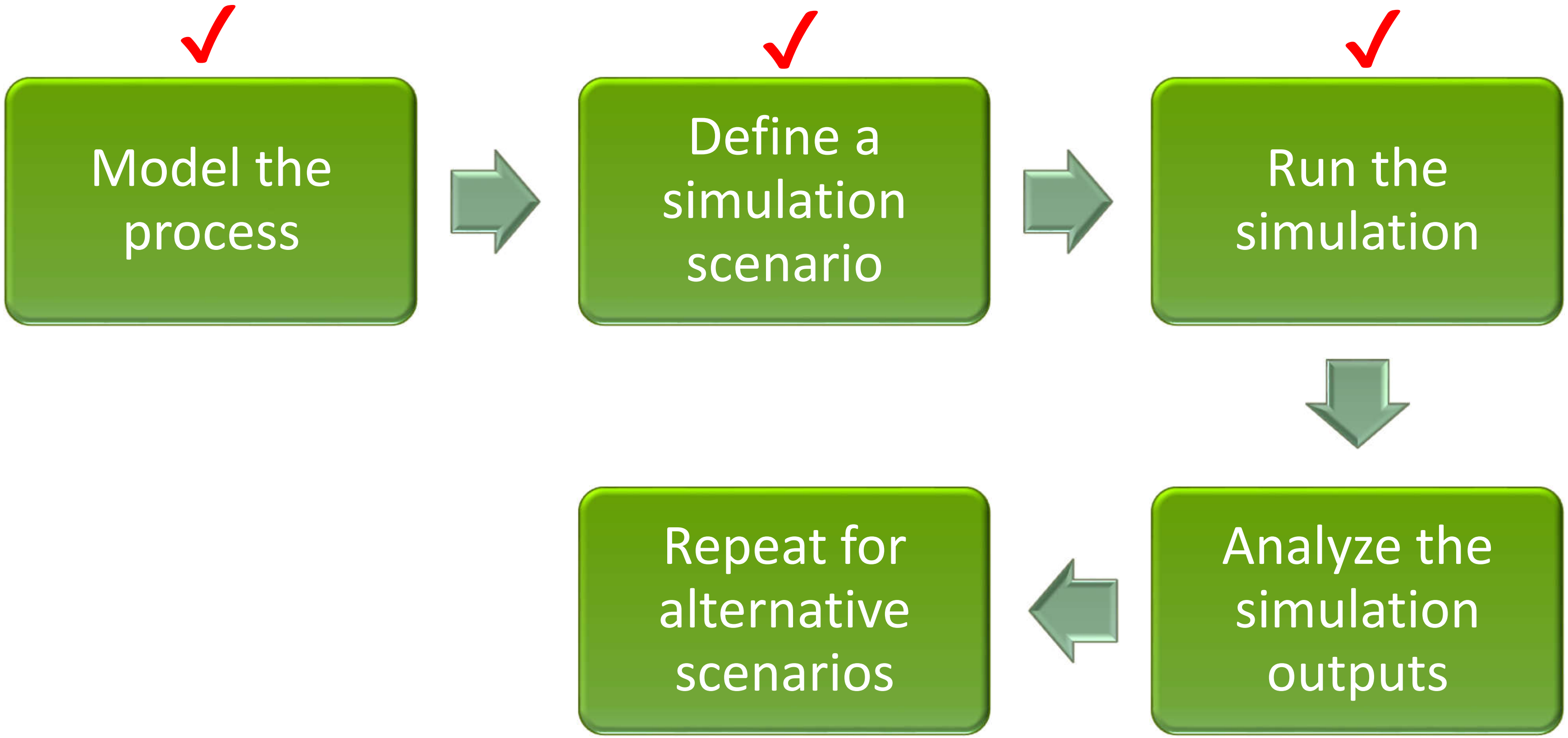
Elements of a simulation model

1. Processing times of activities
   - Fixed value
   - Probability distribution
2. Conditional branching probabilities
3. Arrival rate of process instances and probability distribution
   - Typically exponential distribution with a given mean inter-arrival time
   - Arrival calendar, e.g. Monday-Friday, 9am-5pm, or 24/7
4. Resource pools
5. Assignment of tasks to resource pools

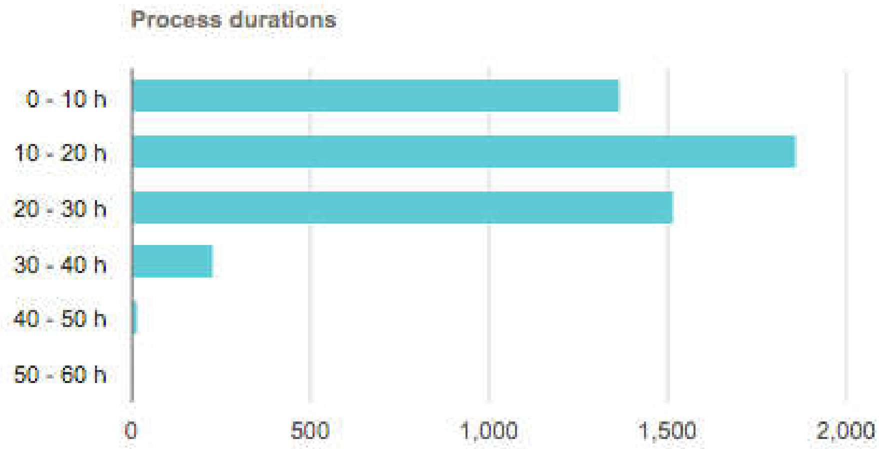
Resource pool assignment

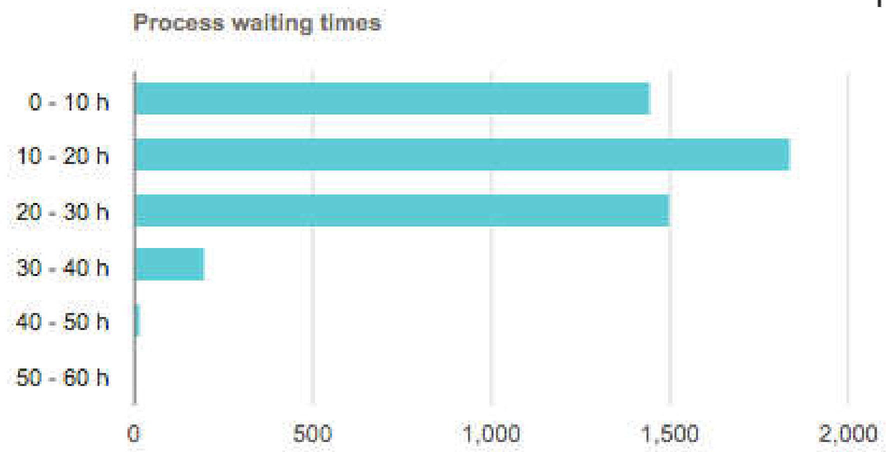


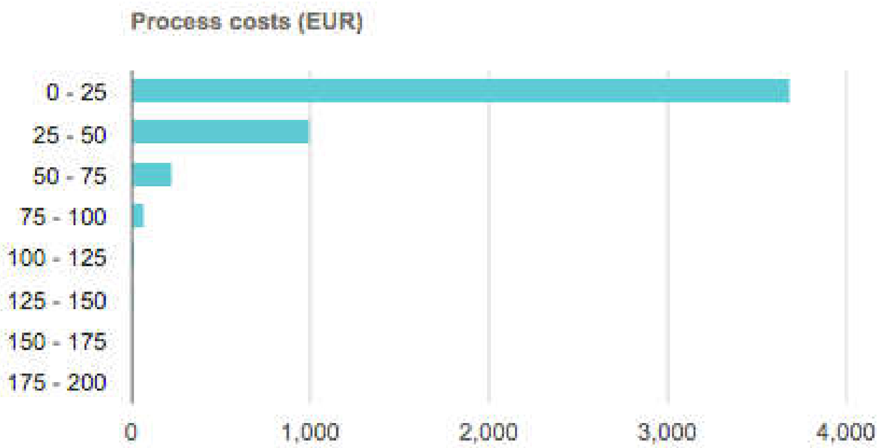
Process Simulation

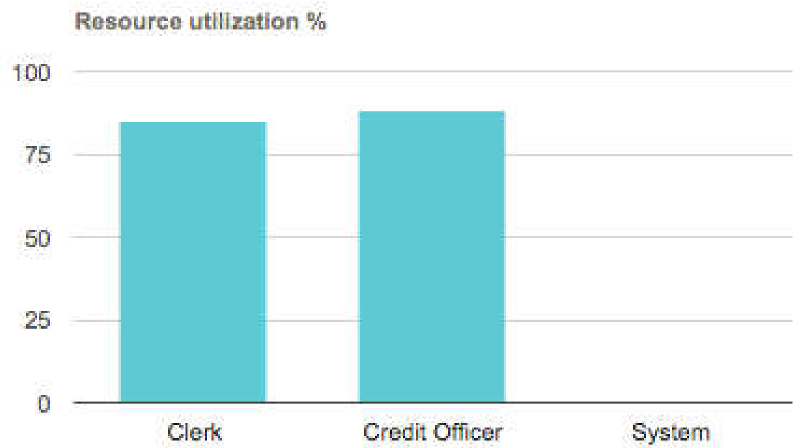


Output: Performance measures & histograms

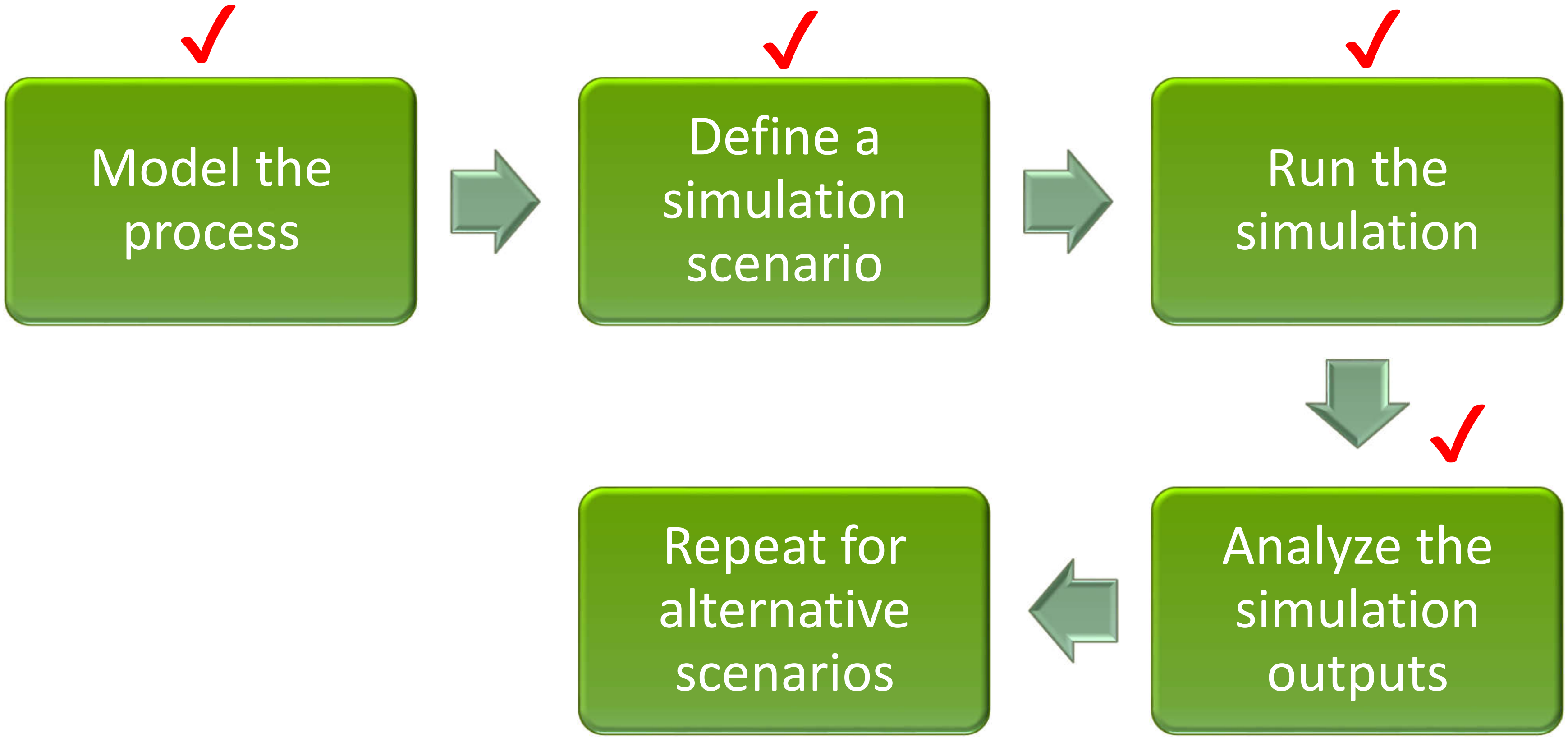








Process Simulation



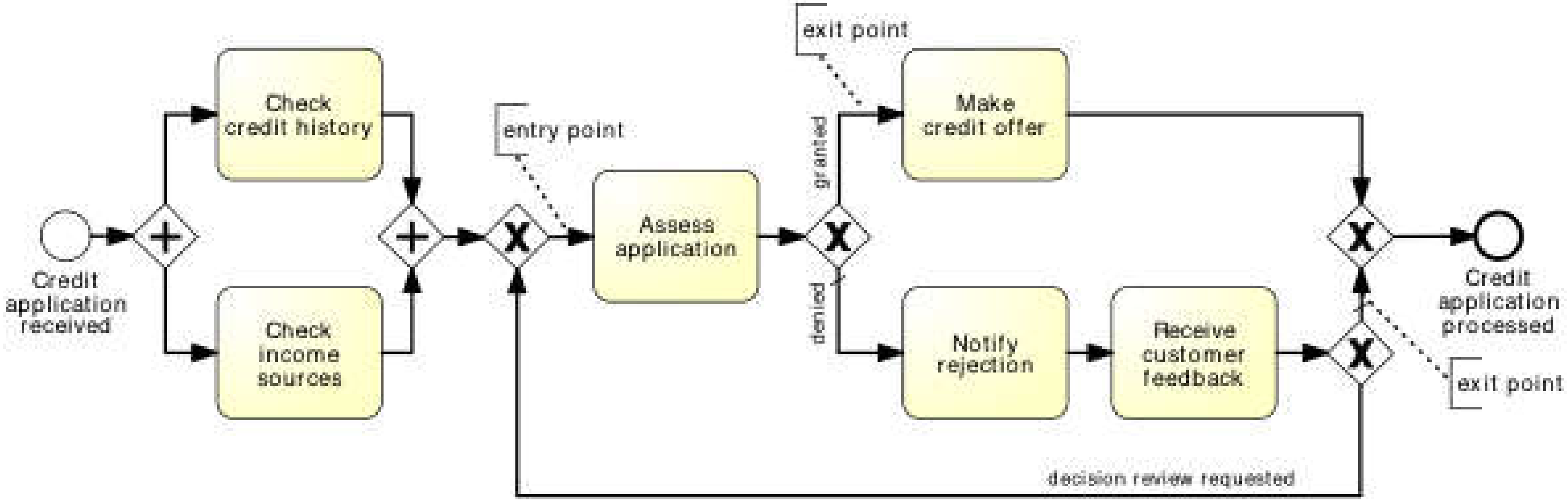
Tools for Process Simulation

- ARIS
- ITP Commerce Process Modeler for Visio
- Logizian
- Oracle BPA
- Progress Savvion Process Modeler
- ProSim
- **Bizagi Process Modeler**
  - **Check tutorial at http://tinyurl.com/bizagisimulation**
- **Signavio + BIMP**
  - **http://bimp.cs.ut.ee/**

BIMP – bimp.cs.ut.ee

- Accepts standard BPMN 2.0 as input
- Simple form-based interface to enter simulation scenario
- Produces KPIs + simulation logs in MXML format
  - Simulation logs can be imported to the ProM process mining tool

BIMP Demo



Credit application received
Check credit history
Check income sources
entry point
Assess application
exit point
granted
denied
Make credit offer
Notify rejection
Receive customer feedback
Credit application processed
exit point
decision review requested

Your turn

- Textbook, Chapter 7, exercise 7.8 (page 240)

Pitfalls of simulation

- Stochasticity
- Data quality
- Simplifying assumptions

Stochasticity

- Problem
  - Simulation results may differ from one run to another
- Solutions
  1. Make the simulation timeframe long enough to cover weekly and seasonal variability, where applicable
  2. Use multiple simulation runs, average results of multiple runs, compute confidence intervals

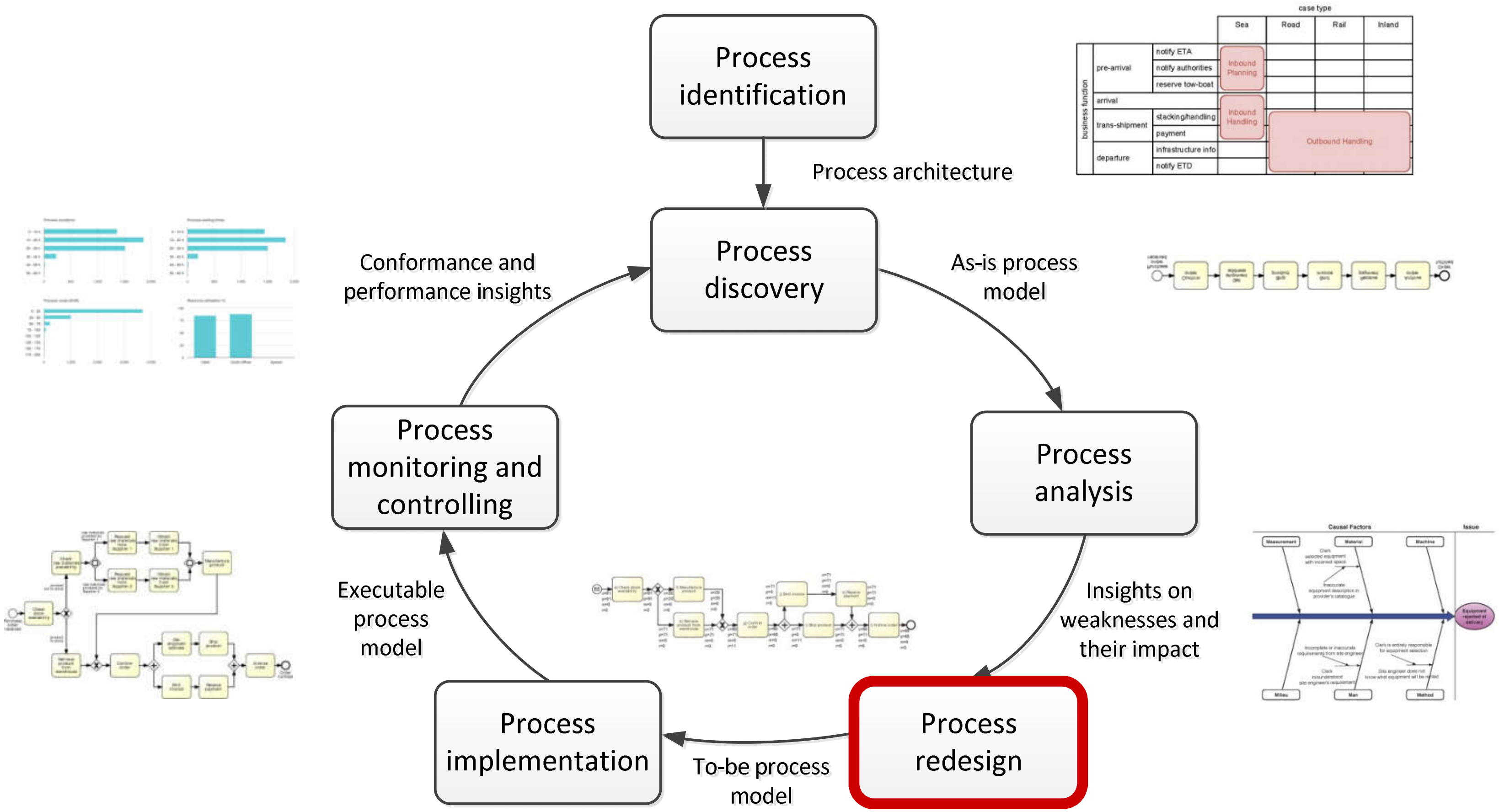
Data quality

- Problem
  - Simulation results are only as trustworthy as the input data
- Solutions:
  1. Rely as little as possible on "guesstimates". Use input analysis where possible:
     - Derive simulation scenario parameters from numbers in the scenario
     - Use statistical tools to check fit the probability distributions
  2. Simulate the "as is" scenario and cross-check results against actual observations

Simulation assumptions

- That the process model is always followed to the letter
  - No deviations
  - No workarounds
- That resources work constantly and non-stop
  - Every day is the same!
  - No tiredness effects
  - No distractions beyond “stochastic” ones

Next week



Process identification
Process architecture
Process discovery
As-is process model
Process analysis
Insights on weaknesses and their impact
Process redesign
To-be process model
Process implementation
Executable process model
Process monitoring and controlling
Conformance and performance insights