MIT 2.852
Manufacturing Systems Analysis
Lectures 2–5

Basic probability, Markov processes, M/M/1 queues

Stanley B. Gershwin

Spring, 2004

Trick Question

I flip a coin 100 times, and it shows heads every time.

Question: What is the probability that it will show heads on the next flip?

Probability and Statistics

Probability $\neq$ *Statistics*

Probability: mathematical theory that describes uncertainty.

Statistics: set of techniques for extracting useful information from data.

Interpretations of probability

Frequency

The probability that the outcome of an experiment is A *is* prob (A)

if the experiment is performed a large number of times and the fraction of times that the observed outcome is A is prob (A).

Interpretations of probability

Parallel universes

The probability that the outcome of an experiment is A *is* prob (A)

if the experiment is performed in each parallel universe and the fraction of universes in which the observed outcome is A is prob (A).

Interpretations of probability

State of belief

The probability that the outcome of an experiment is A *is* prob (A)

if that is the opinion (ie, belief or state of mind) of an observer *before* the experiment is performed.

Interpretations of probability

Abstract measure

The probability that the outcome of an experiment is A
is prob (A)

if prob () satisfies a set of axioms.

Interpretations of probability

Abstract measure

Axioms of probability

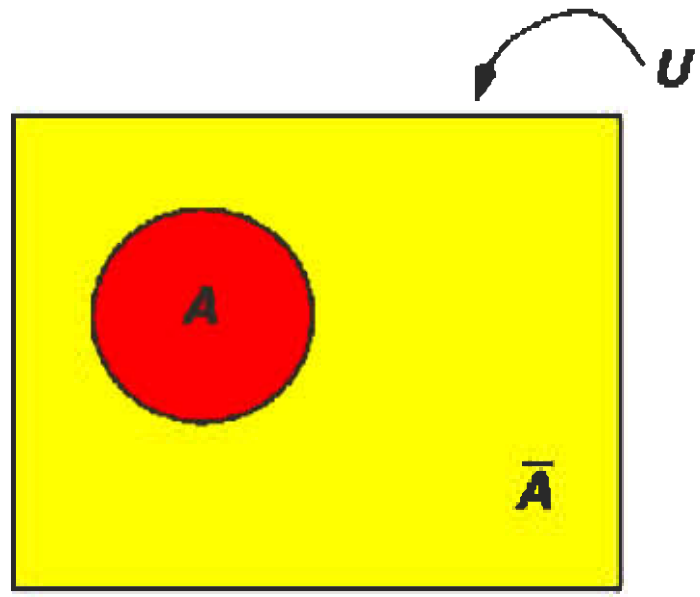
Let U be a set of *samples* . Let E_1, E_2, ... be subsets of U. Let ϕ be the *null set* (the set that has no elements).

- $0 \leq$ prob $(E_i) \leq 1$
- prob $(U) = 1$
- prob $(\phi) = 0$
- If $E_i \cap E_j = \phi$, then
  prob $(E_i \cup E_j) =$ prob $(E_i) +$ prob (E_j)

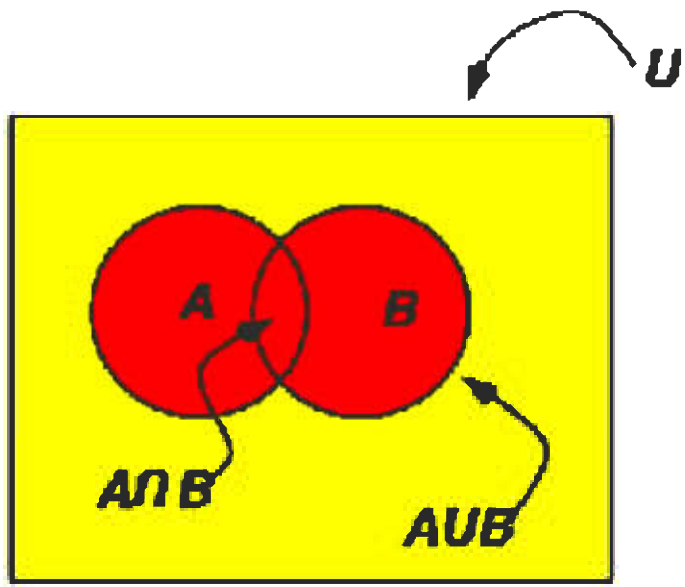
Probability Basics

- Subsets of U are called *events.*

- prob (E) is the *probability* of E.



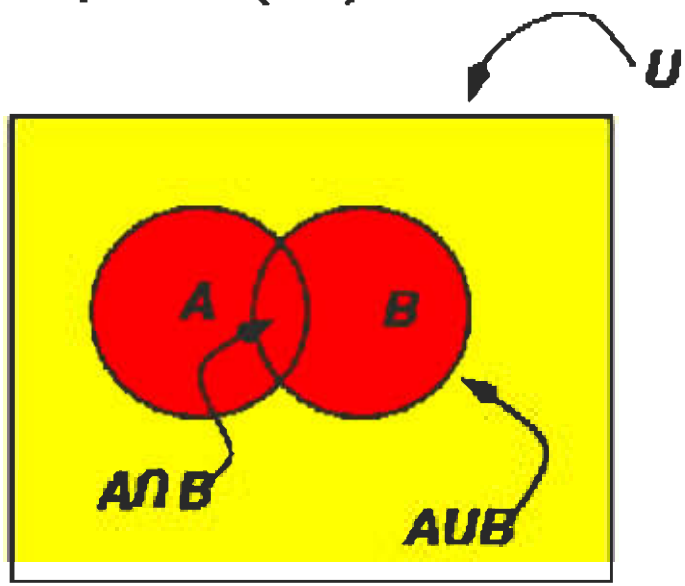
$$\text{prob}\,(\bar{A}) = 1 - \text{prob}\,(A)$$



$$\text{prob}\,(A \cup B) = \text{prob}\,(A) + \text{prob}\,(B) - \text{prob}\,(A \cap B)$$

Conditional Probability

$$\text{prob}\,(A|B) = \frac{\text{prob}\,(A \cap B)}{\text{prob}\,(B)}$$



$$\text{prob}\,(A \cap B) = \text{prob}\,(A|B)\,\text{prob}\,(B).$$

Throw a die.

- $\boldsymbol{A}$ is the event of getting an odd number (1, 3, 5).
- $\boldsymbol{B}$ is the event of getting a number less than or equal to 3 (1, 2, 3).

Then prob $(\boldsymbol{A}) =$ prob $(\boldsymbol{B}) = \mathbf{1/2}$ and
prob $(\boldsymbol{A} \cap \boldsymbol{B}) =$ prob $(\mathbf{1}, \mathbf{3}) = \mathbf{1/3}$.

Also,
prob $(\boldsymbol{A}|\boldsymbol{B}) =$ prob $(\boldsymbol{A} \cap \boldsymbol{B})/$ prob $(\boldsymbol{B}) = \mathbf{2/3}$.

Law of Total Probability

- Let $B = C \cup D$ and assume $C \cap D = \phi$. We have

$$\text{prob}\,(A|C) = \frac{\text{prob}\,(A \cap C)}{\text{prob}\,(C)} \text{ and}$$

$$\text{prob}\,(A|D) = \frac{\text{prob}\,(A \cap D)}{\text{prob}\,(D)}.$$

- Also

$$\text{prob}\,(A \cap B) = \text{prob}\,(A \cap (C \cup D)) =$$
$$\text{prob}\,(A \cap C) + \text{prob}\,(A \cap D) - \text{prob}\,(A \cap (C \cap D)) =$$
$$\text{prob}\,(A \cap C) + \text{prob}\,(A \cap D)$$

Law of Total Probability

- Since

$$\text{prob}\,(A|B)\ \text{prob}\,(B) = \text{prob}\,(A|C)\ \text{prob}\,(C) + \text{prob}\,(A|D)\ \text{prob}\,(D)$$

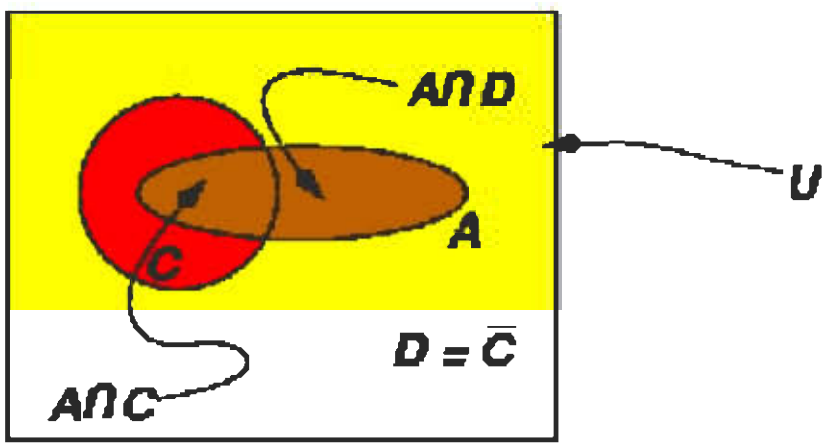
we have

$$\text{prob}\,(A|B) = \text{prob}\,(A|C)\ \text{prob}\,(C|B) + \text{prob}\,(A|D)\ \text{prob}\,(D|B).$$

Law of Total Probability

An important case is when $C \cup D = B = U$, so that $A \cap B = A$. Then

$$\text{prob}\,(A) = \text{prob}\,(A \cap C) + \text{prob}\,(A \cap D) = \text{prob}\,(A|C)\,\text{prob}\,(C) + \text{prob}\,(A|D)\,\text{prob}\,(D).$$



Law of Total Probability

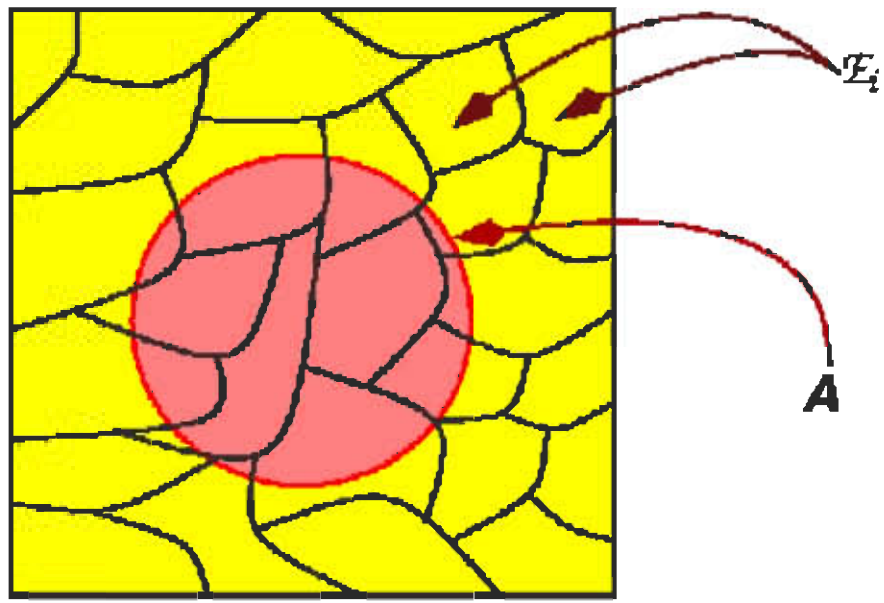
More generally, if A and $\mathcal{E}_1, \ldots \mathcal{E}_k$ are events and

$\mathcal{E}_i$ and $\mathcal{E}_j = \emptyset$, for all $i \neq j$

and

$\bigcup_j \mathcal{E}_j =$ the universal set

(ie, the set of $\mathcal{E}_j$ sets is *mutually exclusive* and *collectively exhaustive*) then ...



Law of Total Probability

$$\sum_j \text{prob}\,(\mathcal{E}_j) = 1$$

and

$$\text{prob}\,(A) = \sum_j \text{prob}\,(A|\mathcal{E}_j)\ \text{prob}\,(\mathcal{E}_j).$$

Law of Total Probability

Some useful generalizations:

$$\text{prob}\,(A|B) = \sum_j \text{prob}\,(A|B \text{ and } \mathcal{E}_j)\ \text{prob}\,(\mathcal{E}_j|B),$$

$$\text{prob}\,(A \text{ and } B) =$$

$$\sum_j \text{prob}\,(A|B \text{ and } \mathcal{E}_j)\ \text{prob}\,(\mathcal{E}_j \text{ and } B).$$

Random Variables

Let V be a vector space. Then a *random variable* X is a mapping (a function) from U to V.

If $\omega \in U$ and $x = X(\omega) \in V$, then X is a random variable.

Dynamic Systems

- t is the time index, a scalar. It can be discrete or continuous.
- $\boldsymbol{X}(t)$ is the state.
  - ⋆ The state can be scalar or vector.
  - ⋆ The state can be discrete or continuous or mixed.
  - ⋆ The state can be deterministic or random.

$\boldsymbol{X}(t)$ is a *stochastic process* if $\boldsymbol{X}(t)$ is a random variable for every t.

Bernoulli

Flip a biased coin. If X^B is *Bernoulli*, then there is a p such that

prob($X^B = 1$) = p.

prob($X^B = 0$) = $1 - p$.

Geometric

The number of Bernoulli random variables X_i^B tested until the first 1 appears is a *geometric* random variable X^g.

$$X^g = \min_i \{X_i^B = 1\}$$

To calculate prob $(X^b = t)$:

- For $t = 1$, we know prob $(X^b = 1) = p$.
  Therefore prob $(X^b > 1) = 1 - p$.

Geometric

- For $t > 1$,

$$\begin{aligned}\text{prob}\,(X^b > t) &= \text{prob}\,(X^b > t | X^b > t-1)\ \text{prob}\,(X^b > t-1) \\ &= (1-p)\ \text{prob}\,(X^b > t-1),\end{aligned}$$

so

$$\text{prob}\,(X^b > t) = (1-p)^t$$

and

$$\text{prob}\,(X^b = t) = (1-p)^{t-1}p$$

Geometric

Alternative view

Consider a two-state system. The system can go from 1 to 0, but not from 0 to 1.

Let p be the conditional probability that the system is in state 0 at time $t+1$, given that it is in state 1 at time t. Then

$$p = \text{ prob } [\alpha(t+1) = 0 | \alpha(t) = 1] \, .$$

Discrete Random Variables

Geometric

Let $\mathbf{p}(\alpha, t)$ be the probability of being in state α at time t.

Then, since

$$\begin{aligned}\mathbf{p}(0, t+1) &= \text{prob } [\alpha(t+1) = 0 | \alpha(t) = 1] \text{ prob } [\alpha(t) = 1] \\ &+ \text{prob } [\alpha(t+1) = 0 | \alpha(t) = 0] \text{ prob } [\alpha(t) = 0],\end{aligned}$$

we have

$$\begin{aligned}\mathbf{p}(0, t+1) &= p\mathbf{p}(1, t) + \mathbf{p}(0, t), \\ \mathbf{p}(1, t+1) &= (1-p)\mathbf{p}(1, t),\end{aligned}$$

Discrete Random Variables

Geometric

and the normalization equation

$$\mathrm{p}(1,t) + \mathrm{p}(0,t) = 1.$$

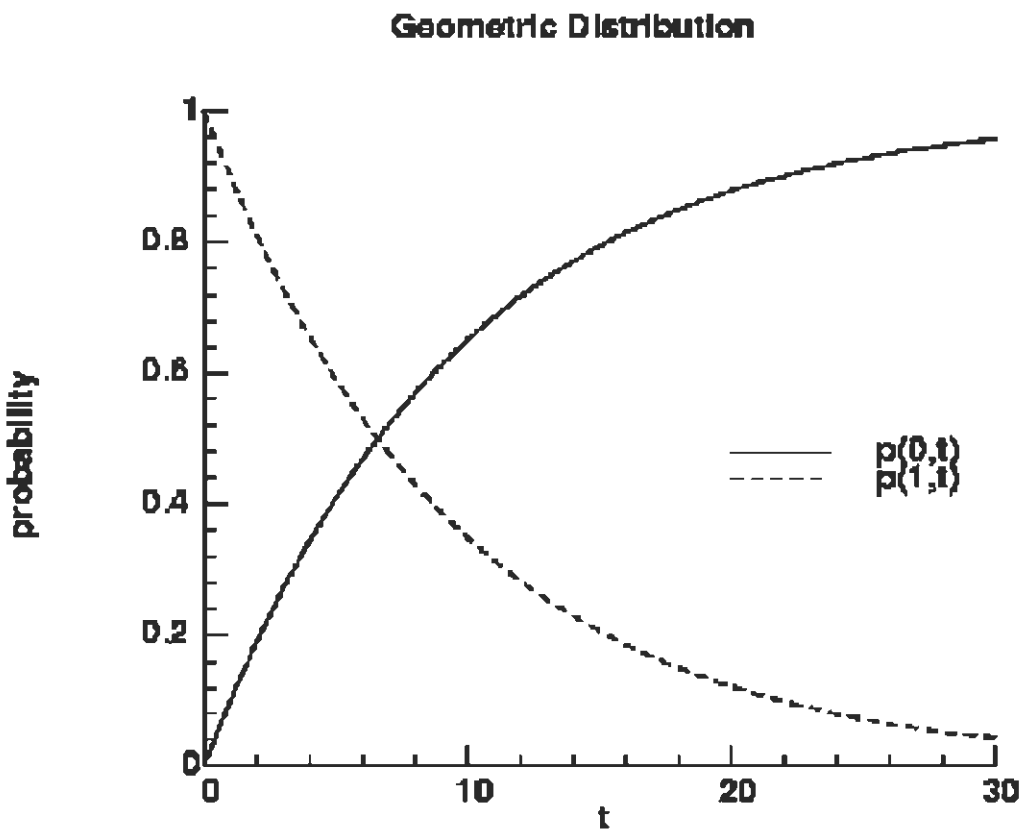
Assume that $\mathrm{p}(1,0) = 1$. Then the solution is

$$\mathrm{p}(0,t) = 1 - (1-p)^t,$$

$$\mathrm{p}(1,t) = (1-p)^t.$$

Discrete Random Variables

Geometric



Discrete Random Variables

Geometric

Recall that once the system makes the transition from 1 to 0 it can never go back. The probability that the transition takes place at time t is

$$\text{prob}\,[\alpha(t) = 0 \text{ and } \alpha(t-1) = 1] = (1-p)^{t-1}p.$$

The time of the transition from 1 to 0 is said to be *geometrically distributed with parameter* p.

The expected transition time is $1/p$. *(Prove it!)*

Geometric

Memorylessness: if T is the transition time,

$$\text{prob}\,(T > t + x | T > x) = \text{prob}\,(T > t).$$

Markov processes

- A *Markov process* is a stochastic process in which the probability of finding X at some value at time $t + \delta t$ depends only on the value of X at time t.

- Or, let $x(s), s \leq t$, be the history of the values of X before time t and let A be a set of possible values of $X(t + \delta t)$. Then

$$\text{prob}\{X(t + \delta t) \in A | X(s) = x(s), s \leq t\} =$$
$$\text{prob}\{X(t + \delta t) \in A | X(t) = x(t)\}$$

Markov processes

- In words: if we know what X was at time t, we don't gain any more useful information about $X(t+\delta t)$ by *also* knowing what X was at any time earlier than t.

Markov processes

States and transitions

Discrete state, discrete time

- States can be numbered 0, 1, 2, 3, ... *(or with multiple indices if that is more convenient).*
- Time can be numbered 0, 1, 2, 3, ... *(or 0, Δ, 2Δ, 3Δ, ... if more convenient).*
- The probability of a transition from j to i in one time unit is often written P_{ij}, where

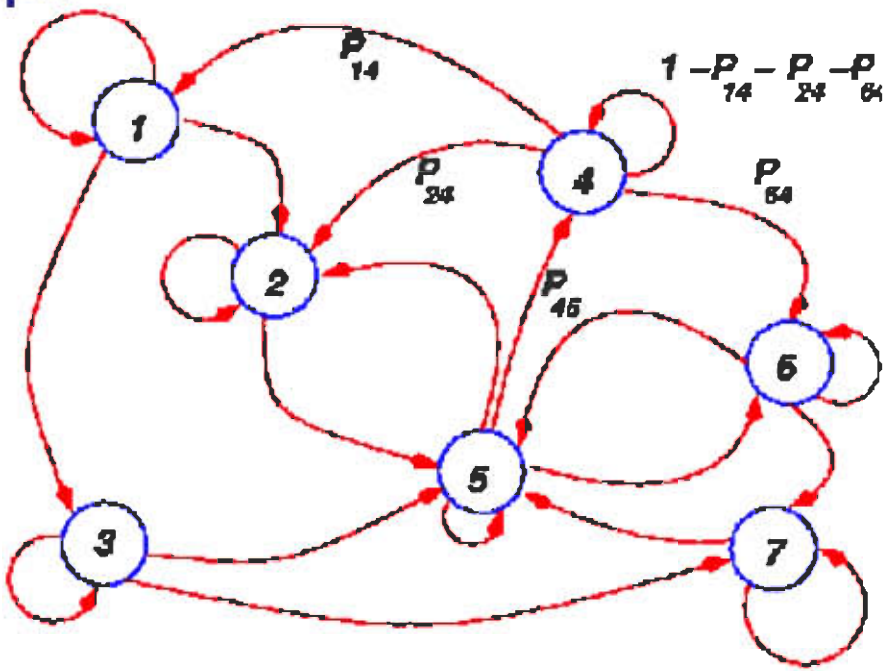
$$P_{ij} = \text{prob}\{X(t+1) = i | X(t) = j\}$$

Markov processes

States and transitions

Discrete state, discrete time

Transition graph



P_{ij} is a probability. Note that $P_{ii} = 1 - \sum_{m, m \neq i} P_{mi}$.

Markov processes

States and transitions

Discrete state, discrete time

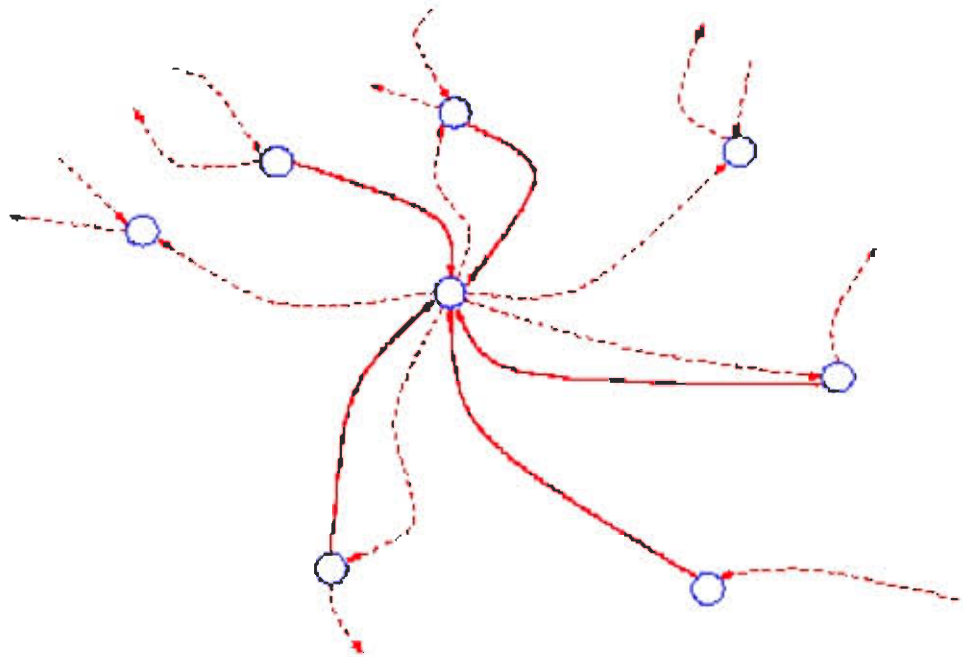
- Define $\mathbf{p}_i(t) = \text{prob}\{X(t) = i\}$
- Transition equations: $\mathbf{p}_i(t+1) = \sum_j P_{ij}\mathbf{p}_j(t)$.
- *Steady state:* $\mathbf{p}_i = \lim_{t\to\infty} \mathbf{p}_i(t)$, if it exists.
- Steady-state transition equations: $\mathbf{p}_i = \sum_j P_{ij}\mathbf{p}_j$.

Markov processes

States and transitions

Discrete state, discrete time

Transition equations are valid for steady-state and non-steady-state conditions.



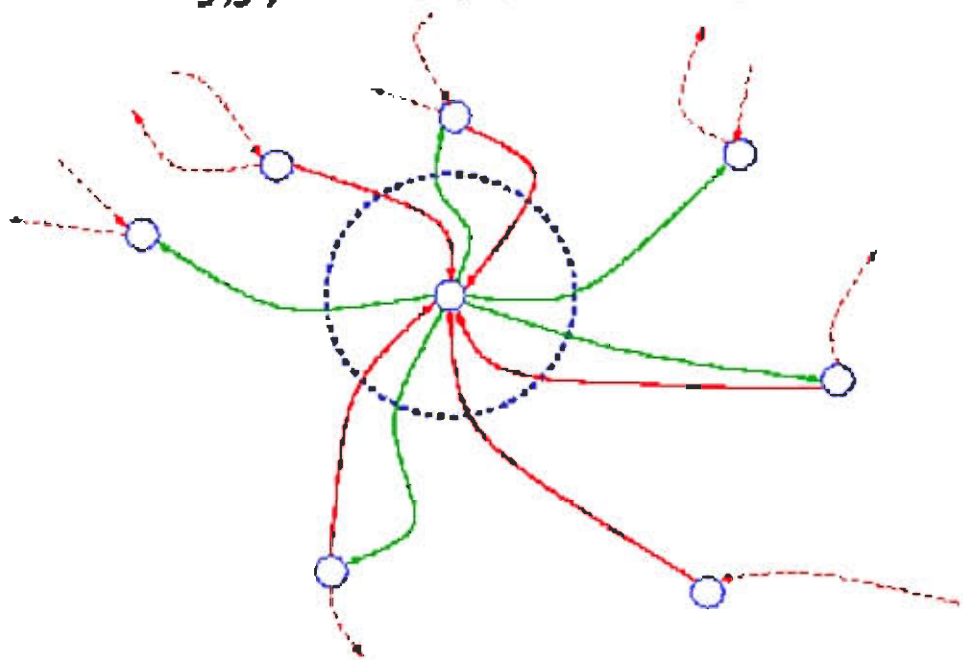
Markov processes

States and transitions

Discrete state, discrete time

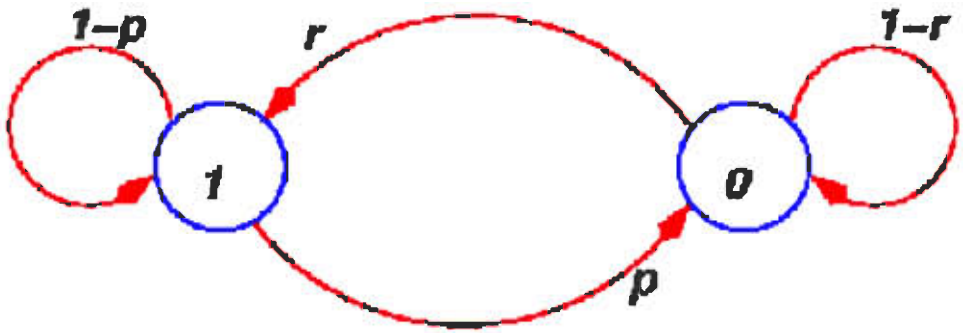
Balance equations — steady-state only. Probability of leaving node i = probability of entering node i.

$$\mathbf{p}_i \sum_{m,m\neq i} P_{mi} = \sum_{j,j\neq i} P_{ij}\mathbf{p}_j$$ *(Prove it!)*



Unreliable machine

1=up; 0=down.



Unreliable machine

The probability distribution satisfies

$$\mathbf{p}(0, t+1) = \mathbf{p}(0, t)(1 - r) + \mathbf{p}(1, t)p,$$

$$\mathbf{p}(1, t+1) = \mathbf{p}(0, t)r + \mathbf{p}(1, t)(1 - p).$$

Unreliable machine

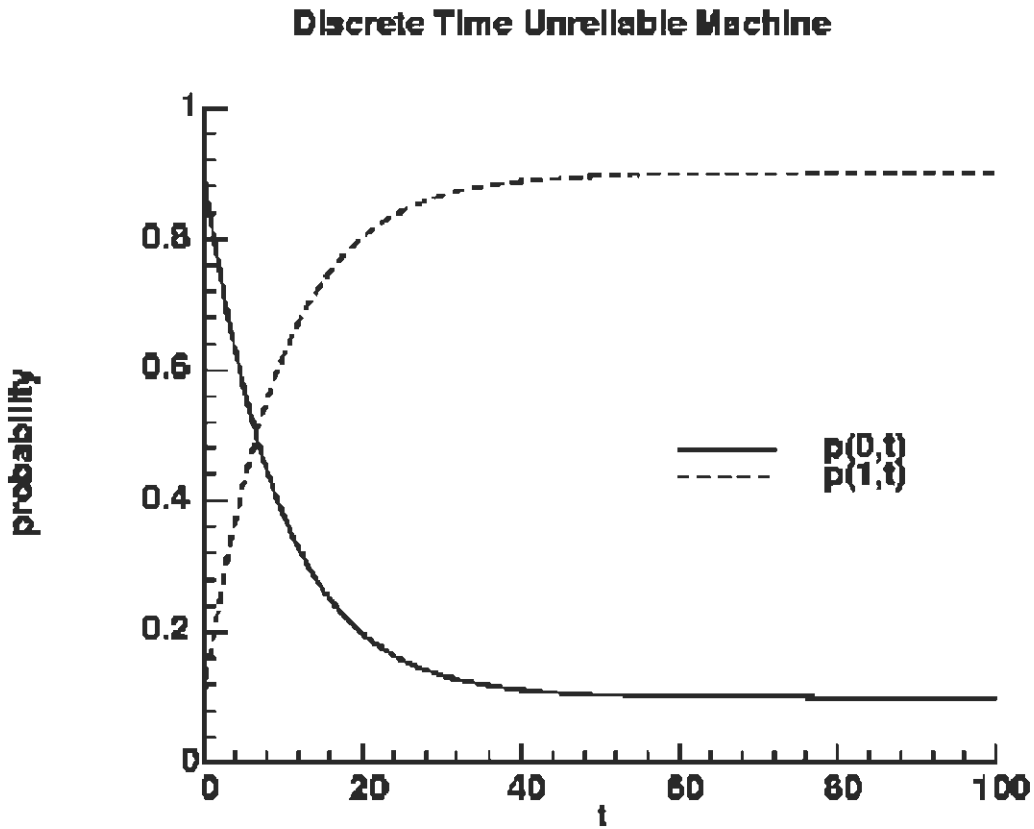
Solution

$$\mathrm{p}(0,t) = \mathrm{p}(0,0)(1-p-r)^t + \frac{p}{r+p}\left[1-(1-p-r)^t\right],$$

$$\mathrm{p}(1,t) = \mathrm{p}(1,0)(1-p-r)^t + \frac{r}{r+p}\left[1-(1-p-r)^t\right].$$

Unreliable machine

Solution



Unreliable machine

Steady-state solution

As $t \to \infty$,

$$\mathrm{p}(0) \to \frac{p}{r+p},$$

$$\mathrm{p}(1) \to \frac{r}{r+p}$$

which is the solution of

$$\mathrm{p}(0) = \mathrm{p}(0)(1-r) + \mathrm{p}(1)p,$$

$$\mathrm{p}(1) = \mathrm{p}(0)r + \mathrm{p}(1)(1-p).$$

If the machine makes one part per time unit when it is operational, the average production rate is

$$\mathrm{p}(1) = \frac{r}{r+p} = \frac{1}{1+\frac{p}{r}}.$$

Markov processes

States and Transitions

Discrete state, continuous time

- States can be numbered 0, 1, 2, 3, ... *(or with multiple indices if that is more convenient).*
- Time is a real number, defined on $(-\infty, \infty)$ or a smaller interval.
- The probability of a transition from j to i during $[t, t+\delta t]$ is approximately $\lambda_{ij}\delta t$, where δt is small, and

$$\lambda_{ij}\delta t = \text{prob}\{X(t+\delta t) = i | X(t) = j\} + o(\delta t).$$

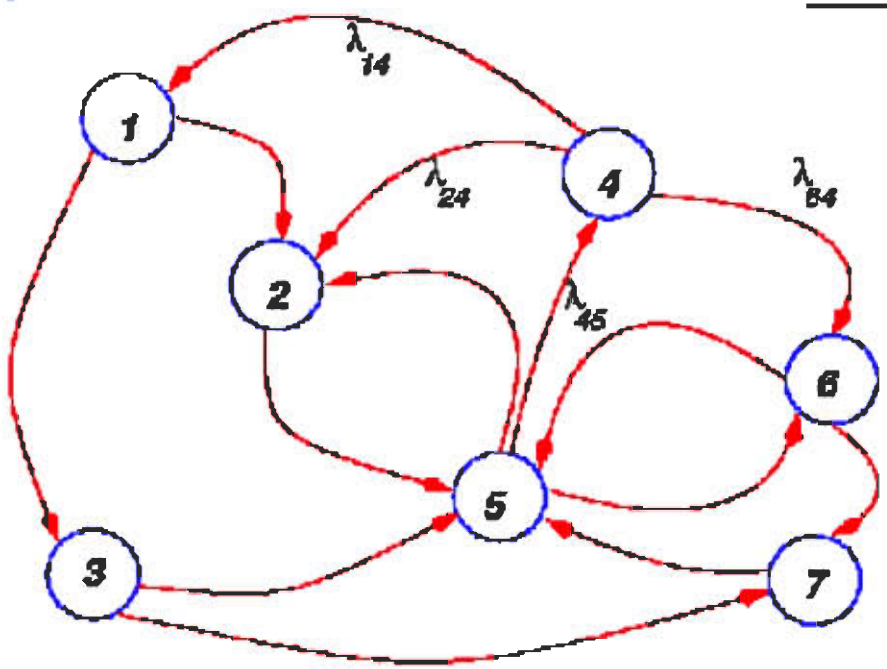
Markov processes

States and Transitions

Discrete state, continuous time

Transition graph

no self loops!!!!



λ_{ij} is a probability *rate*. $\lambda_{ij}\delta t$ is a probability.

Markov processes

States and Transitions

Discrete state, continuous time

- Define $\mathbf{p}_i(t) = \text{prob}\{X(t) = i\}$
- It is convenient to define $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ji}$
- Transition equations: $\frac{d\mathbf{p}_i(t)}{dt} = \sum_j \lambda_{ij} \mathbf{p}_j(t)$.
- *Steady state:* $\mathbf{p}_i = \lim_{t \to \infty} \mathbf{p}_i(t)$, if it exists.
- Steady-state transition equations: $0 = \sum_j \lambda_{ij} \mathbf{p}_j$.
- Steady-state balance equations:
  $\mathbf{p}_i \sum_{m, m \neq i} \lambda_{mi} = \sum_{j, j \neq i} \lambda_{ij} \mathbf{p}_j$

Markov processes

States and Transitions

Discrete state, continuous time

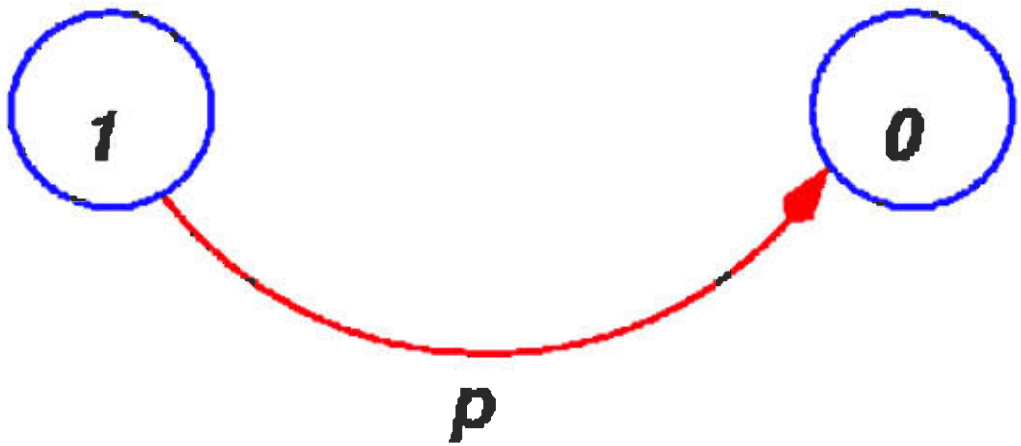
Sources of confusion in continuous time models:

- *Never* Draw self-loops in continuous time markov process graphs.
- *Never* write $1 - P_{14} - P_{24} - P_{64}$. Write
  - $\star 1 - (P_{14} + P_{24} + P_{64})\delta t$, or
  - $\star -(P_{14} + P_{24} + P_{64})$
- $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ji}$ is *NOT* a probability rate and *NOT* a probability. It is *ONLY* a convenient notation.

Markov processes

Exponential

Exponential random variable: the time to move from state 1 to state 0.



$$p\delta t = \text{ prob } [\alpha(t+\delta t) = 0 | \alpha(t) = 1] + o(\delta t).$$

Markov processes

Exponential

$$\mathrm{p}(0, t+\delta t) =$$

$$\text{prob } [\alpha(t+\delta t) = 0 | \alpha(t) = 1] \text{ prob } [\alpha(t) = 1] +$$
$$\text{prob } [\alpha(t+\delta t) = 0 | \alpha(t) = 0] \text{ prob} [\alpha(t) = 0].$$

or

$$\mathrm{p}(0, t+\delta t) = p\delta t \mathrm{p}(1, t) + \mathrm{p}(0, t) + o(\delta t)$$

or

$$\frac{d\mathrm{p}(0, t)}{dt} = p\mathrm{p}(1, t).$$

Markov processes

Exponential

Since $\mathrm{p}(0,t) + \mathrm{p}(1,t) = 1$,

$$\frac{d\mathrm{p}(1,t)}{dt} = -p\mathrm{p}(1,t).$$

If $\mathrm{p}(1,0) = 1$, then

$$\mathrm{p}(1,t) = e^{-pt}$$

and

$$\mathrm{p}(0,t) = 1 - e^{-pt}$$

The probability that the transition takes place in $[t, t + \delta t]$ is

$$\text{prob } [\alpha(t + \delta t) = 0 \text{ and } \alpha(t) = 1] = e^{-pt} p \delta t.$$

The exponential density function is pe^{-pt}.

The time of the transition from 1 to 0 is said to be *exponentially distributed* with rate p. The expected transition time is $1/p$. *(Prove it!)*

Memorylessness: if T is the transition time,

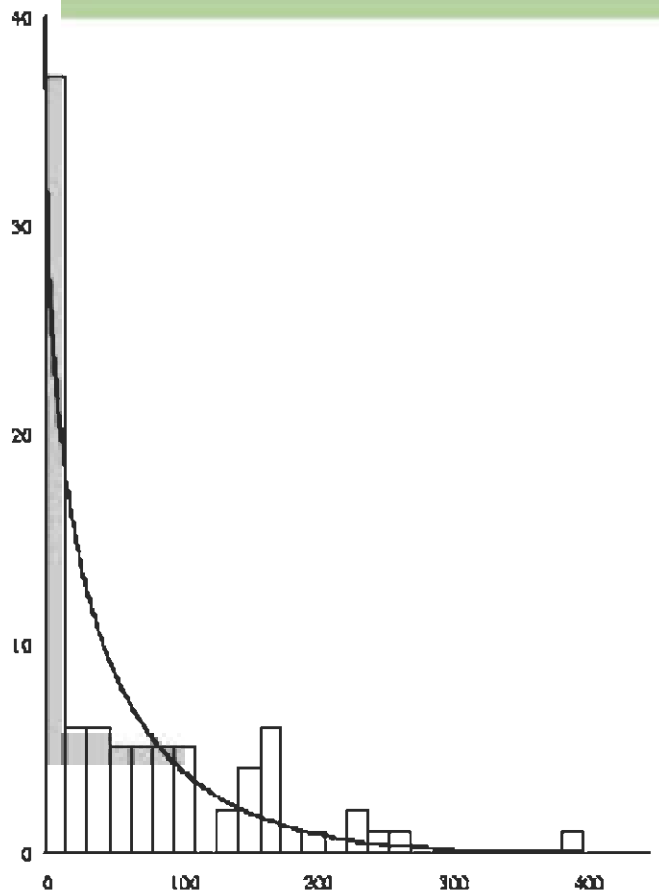
$$\text{prob}\,(T > t + x | T > x) = \text{prob}\,(T > t).$$

Markov processes

Exponential

Density function

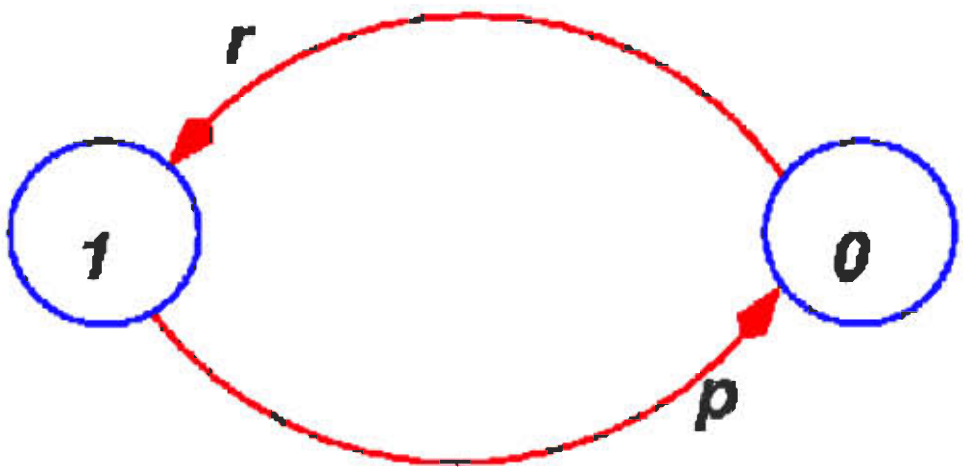
Exponential density function and a small number of actual samples.



Markov processes

Unreliable machine

Continuous time



The probability distribution satisfies

$$\mathrm{p}(0, t+\delta t) = \mathrm{p}(0,t)(1 - r\delta t) + \mathrm{p}(1,t)p\delta t + o(\delta t)$$
$$\mathrm{p}(1, t+\delta t) = \mathrm{p}(0,t)r\delta t + \mathrm{p}(1,t)(1 - p\delta t) + o(\delta t)$$

or

$$\frac{d\mathrm{p}(0,t)}{dt} = -\mathrm{p}(0,t)r + \mathrm{p}(1,t)p$$

$$\frac{d\mathrm{p}(1,t)}{dt} = \mathrm{p}(0,t)r - \mathrm{p}(1,t)p.$$

Unreliable machine

Solution

$$\mathrm{p}(0,t) = \frac{p}{r+p} + \left[\mathrm{p}(0,0) - \frac{p}{r+p}\right] e^{-(r+p)t}$$

$$\mathrm{p}(1,t) = 1 - \mathrm{p}(0,t).$$

As $t \to \infty$,

$$\mathrm{p}(0) \to \frac{p}{r+p},$$

$$\mathrm{p}(1) \to \frac{r}{r+p}$$

If the machine makes μ parts per time unit on the average when it is operational, the overall average production rate is

$$\mu \mathrm{p}(1) = \frac{\mu r}{r+p} = \mu \frac{1}{1+\frac{p}{r}}.$$

Markov processes

The M/M/1 Queue

Consider a queuing system with

- an infinite amount of storage space.
- exponential arrivals:
  - ⋆ If a part arrives at time s, the probability that the next part arrives during the interval $[s+t, s+t+\delta t]$ is $e^{-\lambda t}\lambda\delta t + o(\delta t) \approx \lambda\delta t$. λ is the *arrival rate*.
- exponential service:
  - ⋆ If an operation is completed at time s and the buffer is not empty, the probability that the next operation is completed during the interval $[s+t, s+t+\delta t]$ is $e^{-\mu t}\mu\delta t + o(\delta t) \approx \mu\delta t$. μ is the *service rate*.

The M/M/1 Queue

Let $\mathbf{p}(n,t)$ be the probability that there are n parts in the system at time t. Then,

$$\mathbf{p}(n,t+\delta t) = \mathbf{p}(n-1,t)\lambda\delta t + \mathbf{p}(n+1,t)\mu\delta t$$
$$+\mathbf{p}(n,t)(1-(\lambda\delta t+\mu\delta t)) + o(\delta t)$$
$$\text{for } n > 0$$

and

$$\mathbf{p}(0,t+\delta t) = \mathbf{p}(1,t)\mu\delta t + \mathbf{p}(0,t)(1-\lambda\delta t) + o(\delta t).$$

Or,

$$\frac{d\mathrm{p}(n,t)}{dt} = \mathrm{p}(n-1,t)\lambda + \mathrm{p}(n+1,t)\mu - \mathrm{p}(n,t)(\lambda+\mu),$$
$$n > 0$$
$$\frac{d\mathrm{p}(0,t)}{dt} = \mathrm{p}(1,t)\mu - \mathrm{p}(0,t)\lambda.$$

If a steady state distribution exists, it satisfies

$$0 = \mathrm{p}(n-1)\lambda + \mathrm{p}(n+1)\mu - \mathrm{p}(n)(\lambda+\mu), n > 0$$
$$0 = \mathrm{p}(1)\mu - \mathrm{p}(0)\lambda.$$

Why "if"?

The M/M/1 Queue

Let $\rho = \lambda/\mu$. These equations are satisfied by

$$\mathrm{p}(n) = (1-\rho)\rho^n, n \geq 0$$

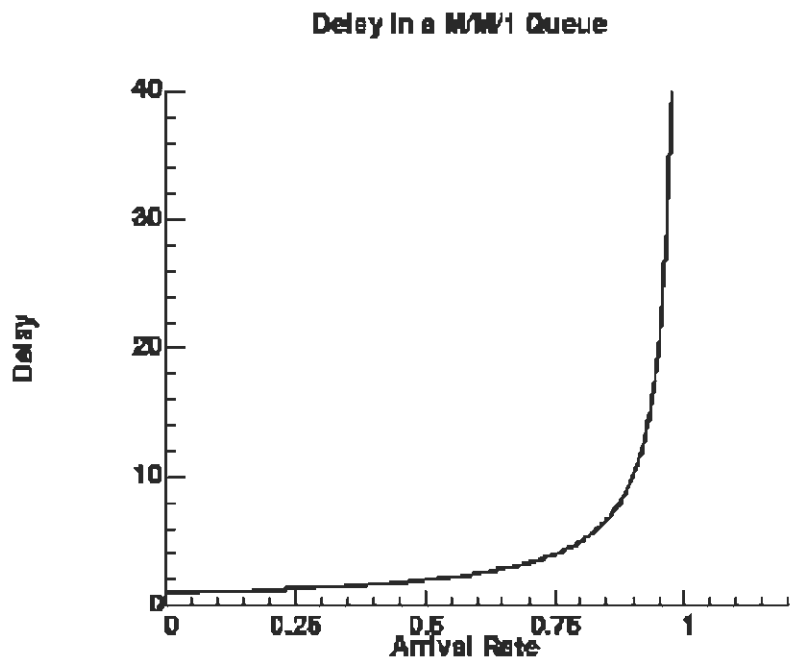
if $\rho < 1$. The average number of parts in the system is

$$\bar{n} = \sum_n n\mathrm{p}(n) = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}.$$

From *Little's law* , the average delay experienced by a part is

$$W = \frac{1}{\mu - \lambda}.$$

The M/M/1 Queue



What happens if $\rho > 1$*?*

Continuous random variables

Philosophical issues

1. Mathematically, continuous and discrete random variables are very different.
2. *Quantitatively* , however, some continuous models are very close to some discrete models.
3. Therefore, which kind of model to use for a given system is a matter of *convenience* .

Continuous random variables

Philosophical issues

Example: The production process for small metal parts (nuts, bolts, washers, etc.) might better be modeled as a continuous flow than a large number of discrete parts.

Continuous random variables

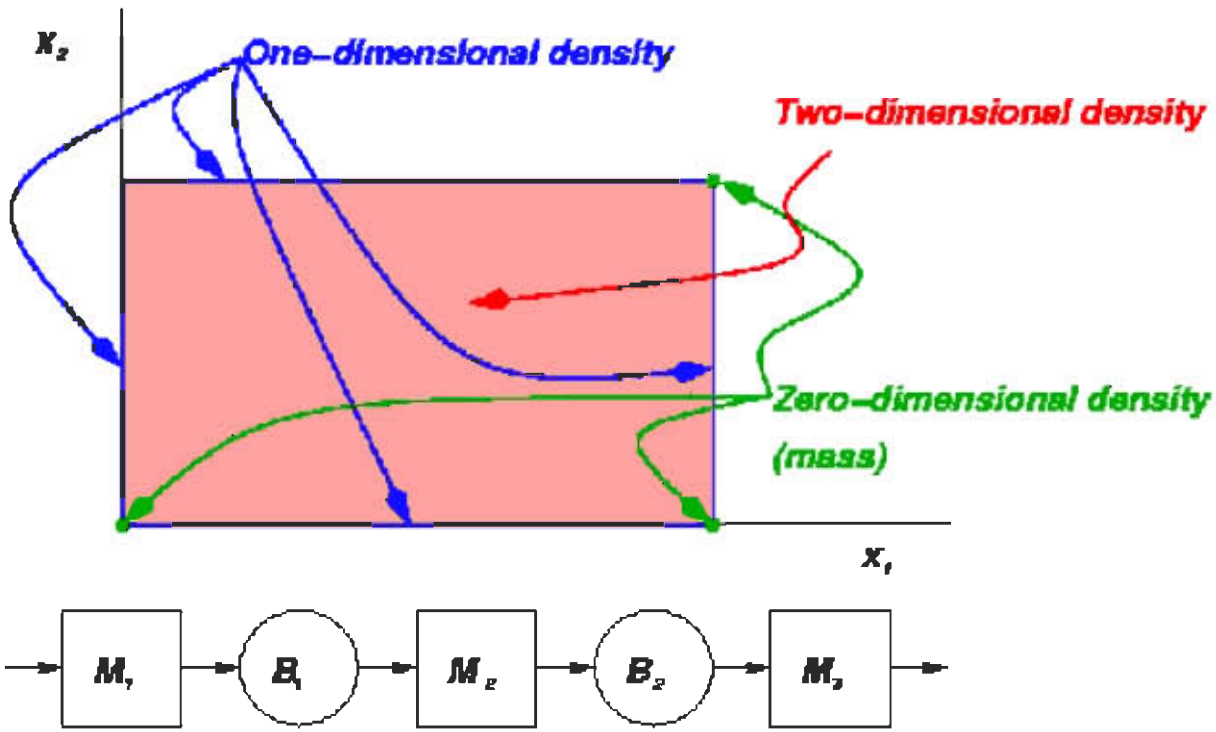
Spaces

- Continuous random variables can be defined
  - ★ in one, two, three, ..., infinite dimensional spaces;
  - ★ in finite or infinite regions of the spaces.
- Continuous random variables can have
  - ★ probability measures with the same dimensionality as the space;
  - ★ lower dimensionality than the space;
  - ★ a mix of dimensions.

Continuous random variables

Spaces

Dimensionality

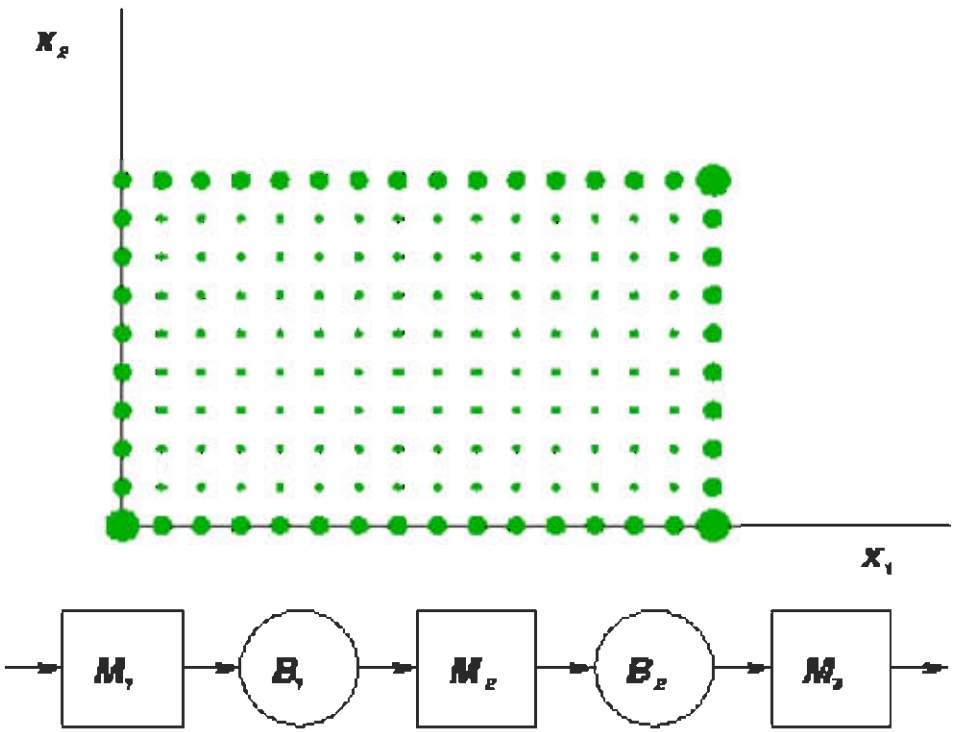


Probability distribution of the amount of material in each of the two buffers.

Continuous random variables

Spaces

Discrete approximation



Probability distribution of the amount of material in each of the two buffers.