Modeling and Verification of Probabilistic Systems

Lecture 2: Discrete-Time Markov Chains

Joost-Pieter Katoen

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

http://www-i2.informatik.rwth-aachen.de/i2/mvps11/

April 12, 2011

Overview

Geometric distribution

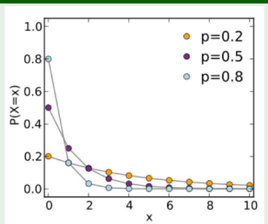
Geometric distribution

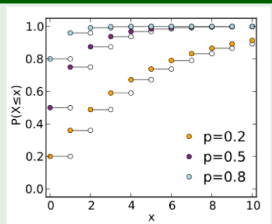
Let X be a discrete random variable, natural $k > 0$ and $0 < p \leqslant 1$. The mass function of a *geometric distribution* is given by:

$$\Pr\{ X = k \} = (1-p)^{k-1} \cdot p$$

We have $E[X] = \frac{1}{p}$ and $Var[X] = \frac{1-p}{p^2}$ and cdf $\Pr\{ X \leqslant k \} = 1 - (1-p)^k$.

Geometric distributions and their cdf's





Memoryless property

Theorem

1. For any random variable X with a geometric distribution:

$$\Pr\{X = k + m \mid X > m\} \;=\; \Pr\{X = k\} \quad \text{for any} \quad m \in T, k \geqslant 1$$

This is called the memoryless property, and X is a memoryless r.v..

2. Any discrete random variable which is memoryless is geometrically distributed.

Proof:

On the black board.

Joint distribution function

Joint distribution function

The *joint* distribution function of stochastic process $X = \{ X_t \mid t \in T \}$ is given for n, $t_1, \ldots, t_n \in T$ and $d_1, \ldots, d_n$ by:

$$F_X(d_1, \ldots, d_n; t_1, \ldots, t_n) = \Pr\{ X(t_1) \leqslant d_1, \ldots, X(t_n) \leqslant d_n \}$$

The shape of F_X depends on the stochastic dependency between $X(t_i)$.

Stochastic independence

Random variables X_i on probability space $\mathcal{P}$ are *independent* if:

$$F_X(d_1, \ldots, d_n; t_1, \ldots, t_n) \;=\; \prod_{i=1}^{n} F_X(d_i; t_i) \;=\; \prod_{i=1}^{n} \Pr\{ X(t_i) \leqslant d_i \}.$$

A renewal process is a discrete-time stochastic process where $X(t_1), X(t_2), \ldots$ are independent, identically distributed, non-negative random variables.

The next state of the stochastic process only depends on the current state, and

Markov property

Markov process

A discrete-time stochastic process $\{ X(t) \mid t \in T \}$ over state space $\{ d_0, d_1, \ldots \}$ is a *Markov process* if for any $t_0 < t_1 < \ldots < t_n < t_{n+1}$:

$$\Pr\{ X(t_{n+1}) = d_{n+1} \mid X(t_0) = d_0, X(t_1) = d_1, \ldots, X(t_n) = d_n \}$$
$$=$$
$$\Pr\{ X(t_{n+1}) = d_{n+1} \mid X(t_n) = d_n \}$$

The distribution of $X(t_{n+1})$, given the values $X(t_0)$ through $X(t_n)$, only depends on the current state $X(t_n)$.

The conditional probability distribution of future states of a Markov process only depends on the current state and not on its further history.

Invariance to time-shifts

Time homogeneity

Markov process $\{ X(t) \mid t \in T \}$ is *time-homogeneous* iff for any $t' < t$:

$$\Pr\{ X(t) = d \mid X(t') = d' \} \;=\; \Pr\{ X(t - t') = d \mid X(0) = d' \}.$$

A time-homogeneous stochastic process is invariant to time shifts.

Discrete-time Markov chain

A *discrete-time Markov chain* (DTMC) is a time-homogeneous Markov process with discrete parameter T and discrete state space.

Discrete-time Markov chain

Discrete-time Markov chain

A *discrete-time Markov chain* (DTMC) is a time-homogeneous Markov process with discrete parameter T and discrete state space S.

Transition probabilities

The *(one-step) transition probability* from $s \in S$ to $s' \in S$ at epoch $n \in \mathbb{N}$ is given by:

$$p^{(n)}(s, s') \;=\; \Pr\{ X_{n+1} = s' \mid X_n = s \} \;=\; \Pr\{ X_1 = s' \mid X_0 = s \}$$

where the last equality is due to time-homogeneity.

Since $p^{(n)}(\cdot) = p^{(k)}(\cdot)$, the superscript (n) is omitted, and we write $p(\cdot)$.

Transition probability matrix

Discrete-time Markov chain

A *discrete-time Markov chain* (DTMC) is a time-homogeneous Markov process with discrete parameter T and discrete state space S.

Transition probability matrix

Let $\mathbf{P}$ be a function with $\mathbf{P}(s_i, s_j) = p(s_i, s_j)$. For finite state space S, function $\mathbf{P}$ is called the *transition probability matrix* of the DTMC with state space S.

Properties

1. $\mathbf{P}$ is a (right) *stochastic* matrix, i.e., it is a square matrix, all its elements are in $[0, 1]$, and each row sum equals one.
2. Matrix $\mathbf{P}$ has an eigenvalue of one, and all its eigenvalues are at most one.
3. For all $n \in \mathbb{N}$, $\mathbf{P}^n$ is a stochastic matrix.

Example: roulette in Monte Carlo, 1913

DTMCs — A transition system perspective

Discrete-time Markov chain

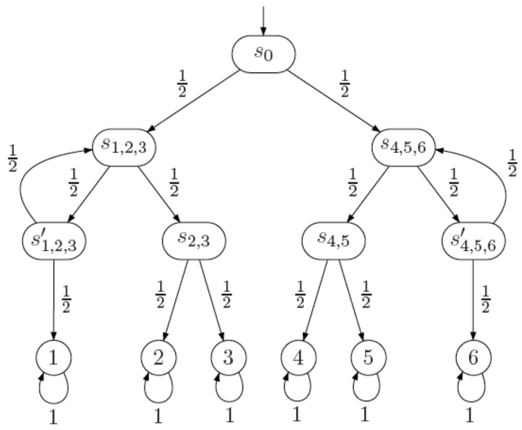
A DTMC $\mathcal{D}$ is a tuple $(S, \mathbf{P}, \iota_{init}, AP, L)$ with:

- S is a countable nonempty set of states
- $\mathbf{P} : S \times S \to [0, 1]$, transition probability function s.t. $\sum_{s'} \mathbf{P}(s, s') = 1$
- $\iota_{init} : S \to [0, 1]$, the initial distribution with $\sum_{s \in S} \iota_{init}(s) = 1$
- AP is a set of atomic propositions.
- $L : S \to 2^{AP}$, the labeling function, assigning to state s, the set $L(s)$ of atomic propositions that are valid in s.

Initial states

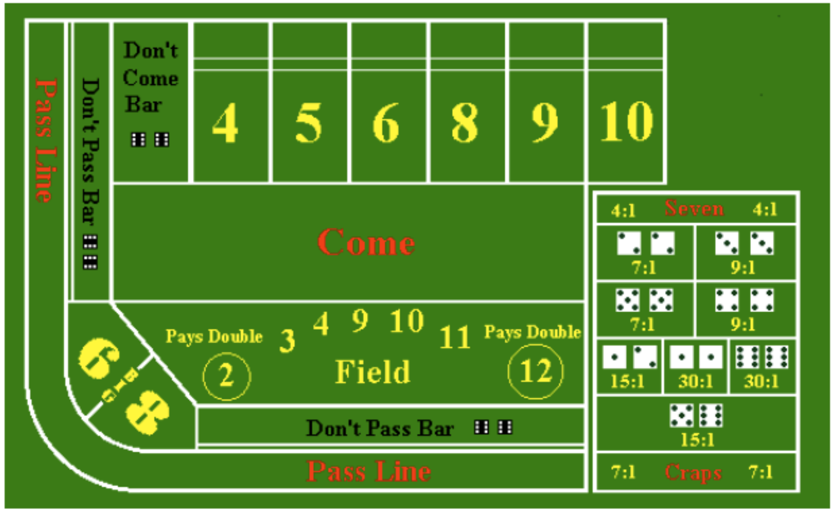
- $\iota_{init}(s)$ is the probability that DTMC $\mathcal{D}$ starts in state s
- the set $\{ s \in S \mid \iota_{init}(s) > 0 \}$ are the possible initial states.

Simulating a die by a fair coin [Knuth & Yao]



Heads = "go left"; tails = "go right". Does this DTMC adequately model a fair six-sided die?

Craps

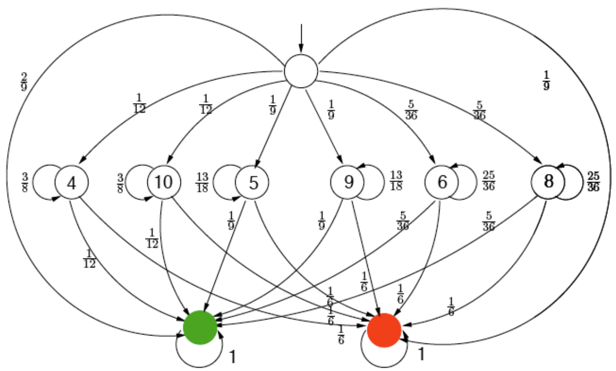


Craps

- Roll two dice and bet
- Come-out roll ("pass line" wager):
  - outcome 7 or 11: win
  - outcome 2, 3, or 12: lose ("craps")
  - any other outcome: roll again (outcome is "point")
- Repeat until 7 or the "point" is thrown:
  - outcome 7: lose ("seven-out")
  - outcome the point: win
  - any other outcome: roll again

A DTMC model of Craps

- Come-out roll:
  - 7 or 11: win
  - 2, 3, or 12: lose
  - else: roll again
- Next roll(s):
  - 7: lose
  - point: win
  - else: roll again



State residence time distribution

Let T_s be the number of epochs of DTMC $\mathcal{D}$ to stay in state s:

$$\begin{aligned}
\Pr\{\, T_s = 1 \,\} &= 1 - \mathbf{P}(s,s) \\
\Pr\{\, T_s = 2 \,\} &= \mathbf{P}(s,s) \cdot (1 - \mathbf{P}(s,s)) \\
\ldots\ldots &\ \ldots\ \ldots\ldots \\
\Pr\{\, T_s = n \,\} &= \mathbf{P}(s,s)^{n-1} \cdot (1 - \mathbf{P}(s,s))
\end{aligned}$$

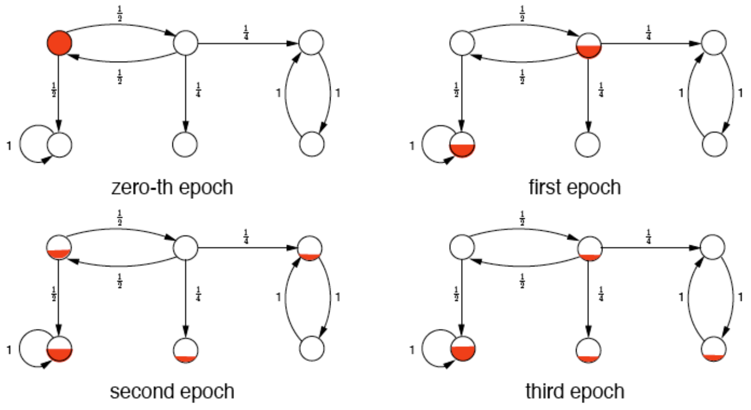
So, the state residence times in a DTMC obey a *geometric* distribution.

The expected number of time steps to stay in state s equals $E[T_s] = \frac{1}{1-\mathbf{P}(s,s)}$.

The variance of the residence time distribution is $Var[T_s] = \frac{\mathbf{P}(s,s)}{(1-\mathbf{P}(s,s))^2}$.

Recall that the geometric distribution is the only discrete probability distribution that possesses the memoryless property.

Evolution of an example DTMC



We want to determine $p_{s,s'}(n) = \Pr\{ X(n) = s' \mid X(0) = s \}$ for $n \in \mathbb{N}$.

Determining n-step transition probabilities

n-step transition probabilities

The probability to move from s to s' in $n \in \mathbb{N}$ steps is inductively defined:

$$p_{s,s'}(0) \;=\; 1 \quad \text{if } s = s', \quad \text{and 0 otherwise,}$$

$p_{s,s'}(1) = \mathbf{P}(s, s')$, and for $n > 1$ by the Chapman-Kolmogorov equation:

$$p_{s,s'}(n) = \sum_{s''} p_{s,s''}(l) \cdot p_{s'',s'}(n{-}l) \quad \text{for all } 0 < l < n$$

Proof: see blackboard.

For $l = 1$ and $n > 0$ we obtain: $p_{s,s'}(n) = \sum_{s''} p_{s,s''}(1) \cdot p_{s'',s'}(n{-}1)$

$\mathbf{P}^{(n)} = \mathbf{P}^{(1)} \cdot \mathbf{P}^{(n-1)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)}$ is the n-step transition probability matrix

Repeating this scheme: $\mathbf{P}^{(n)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)} = \ldots = \mathbf{P}^{n-1} \cdot \mathbf{P}^{(1)} = \mathbf{P}^{n}$.

Transient probability distribution

Transient distribution

$\mathbf{P}^n(s,t)$ equals the probability of being in state t after n steps given that the computation starts in s.

The probability of DTMC $\mathcal{D}$ being in state t after exactly n transitions is:

$$\Theta^{\mathcal{D}}_n(t) \;=\; \sum_{s \in S} \iota_{init}(s) \cdot \mathbf{P}^n(s,t)$$

$\Theta^{\mathcal{D}}_n(t)$ is called the *transient state probability* at epoch n for state t. The function $\Theta^{\mathcal{D}}_n$ is the *transient state distribution* at epoch n of DTMC $\mathcal{D}$.

When considering $\Theta^{\mathcal{D}}_n$ as vector $(\Theta^{\mathcal{D}}_n)_{t \in S}$ we have:

$$\Theta^{\mathcal{D}}_n \;=\; \iota_{init} \cdot \underbrace{\mathbf{P} \cdot \mathbf{P} \cdot \ldots \cdot \mathbf{P}}_{n \text{ times}} \;=\; \iota_{init} \cdot \mathbf{P}^n.$$

Transient probability distribution: example

Paths in a DTMC

State graph

The *state graph* of DTMC $\mathcal{D}$ is a digraph $G = (V, E)$ with V are the states of $\mathcal{D}$, and $(s, s') \in E$ iff $\mathbf{P}(s, s') > 0$.

Paths

Paths in $\mathcal{D}$ are maximal (i.e., infinite) paths in its state graph. Thus, a path is an infinite sequence of states $s_0 s_1 s_2 \ldots \ldots$ with $\mathbf{P}(s_i, s_{i+1}) > 0$ for all i.

Let $Paths(\mathcal{D})$ denote the set of paths in $\mathcal{D}$, and $Paths^*(\mathcal{D})$ the set of finite prefixes thereof.

Direct successors and predecessors

$Post(s) = \{ s' \in S \mid \mathbf{P}(s, s') > 0 \}$ and $Pre(s) = \{ s' \in S \mid \mathbf{P}(s', s) > 0 \}$ are the set of direct successors and predecessors of s respectively. $Post^*(s)$ and $Pre^*(s)$ are the reflexive and transitive closure of $Post$ and Pre.

Paths and probabilities

To reason quantitatively about the behavior of a DTMC, we need to define a probability space over its paths.

Intuition

For a given state s in DTMC $\mathcal{D}$:

- Sample space := set of all infinite paths starting in s
- Events := sets of infinite paths starting in s
- Basic events := cylinder sets
- Cylinder set of finite path $\hat{\pi}$:= set of all infinite continuations of $\hat{\pi}$

Probability measure on DTMCs

Cylinder set

The *cylinder set* of finite path $\hat{\pi} = s_0\, s_1 \ldots s_n \in Paths^*(\mathcal{D})$ is defined by:

$$Cyl(\hat{\pi}) \;=\; \{ \pi \in Paths(\mathcal{D}) \mid \hat{\pi} \text{ is a prefix of } \pi \}$$

The cylinder set spanned by finite path $\hat{\pi}$ thus consists of all infinite paths that have prefix $\hat{\pi}$. Cylinder sets serve as basic events of the smallest σ-algebra on $Paths(\mathcal{D})$.

σ-algebra of a DTMC

The σ-algebra associated with DTMC $\mathcal{D}$ is the smallest σ-algebra that contains all cylinder sets $Cyl(\hat{\pi})$ where $\hat{\pi}$ ranges over all finite path fragments in $\mathcal{D}$.

Probability measure on DTMCs

Cylinder set

The cylinder set of finite path $\hat{\pi} = s_0\, s_1 \ldots s_n \in Paths^*(\mathcal{D})$ is defined by:

$$Cyl(\hat{\pi}) \;=\; \{ \pi \in Paths(\mathcal{D}) \mid \hat{\pi} \text{ is a prefix of } \pi \}$$

Probability measure

Pr is the unique *probability measure* on the σ-algebra on *$Paths(\mathcal{D})$* defined by:

$$\Pr(Cyl(s_0 \ldots s_n)) \;=\; \iota_{init}(s_0) \cdot \mathbf{P}(s_0\, s_1 \ldots s_n)$$

where $\mathbf{P}(s_0\, s_1 \ldots s_n) \;=\; \prod_{0 \leqslant i < n} \mathbf{P}(s_i, s_{i+1})$ for $n > 0$ and $\mathbf{P}(s_0) = 1$.

Example