Primer: Markov chains

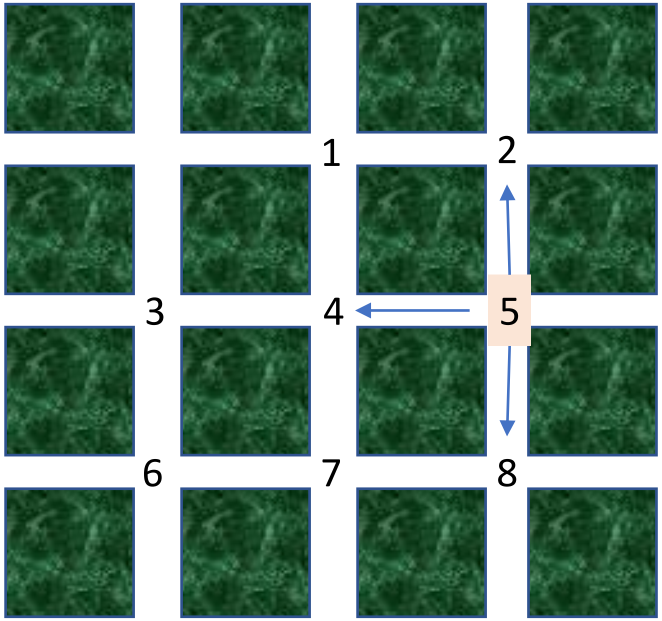
Lecture 7.2

by Marina Barsky

See sample code in *[markov.py]*

Traffic officer assignment

- One traffic officer is assigned to control intersections: 1 - 8
- He is instructed to remain at a given intersection for an hour, and then either remain at the same intersection, or move to a neighboring intersection
- To avoid establishing a pattern, he is instructed to choose his new intersection at random, with each possible move - equally likely



For example, if he starts at intersection 5, his next intersection could be 2,4,5,8 – all with equal probability ¼

Every day he starts at the location where he stopped before

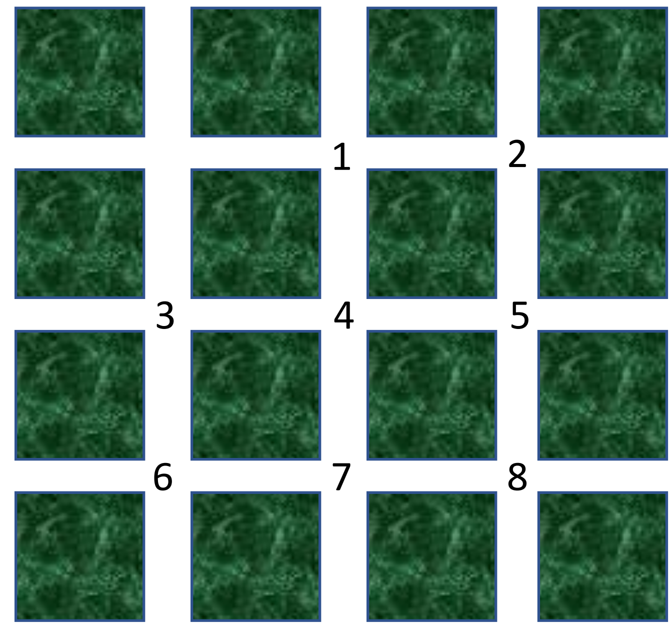
Markov chains

- The system can be in a finite number of states
- The transition from state to state is not predetermined, but rather specified in terms of *probabilities* which depend on the previous history of the system. Such a system is called a *stochastic* system

- If the transition probabilities depend only on the immediate history of the system – for example the state at current observation depends only on the state in the preceding observation – then *the process of transitions from state to state* is called a ***Markov process*** or a ***Markov chain***

Markov model for traffic officer

		Old intersection							
		1	2	3	4	5	6	7	8
New intersection	1	1/3	1/3	0	1/5	0	0	0	0
	2	1/3	1/3	0	0	1/4	0	0	0
	3	0	0	1/3	1/5	0	1/3	0	0
	4	1/3	0	1/3	1/5	1/4	0	1/4	0
	5	0	1/3	0	1/5	1/4	0	0	1/3
	6	0	0	1/3	0	0	1/3	1/4	0
	7	0	0	0	1/5	0	1/3	1/4	1/3
	8	0	0	0	0	1/4	0	1/4	1/3



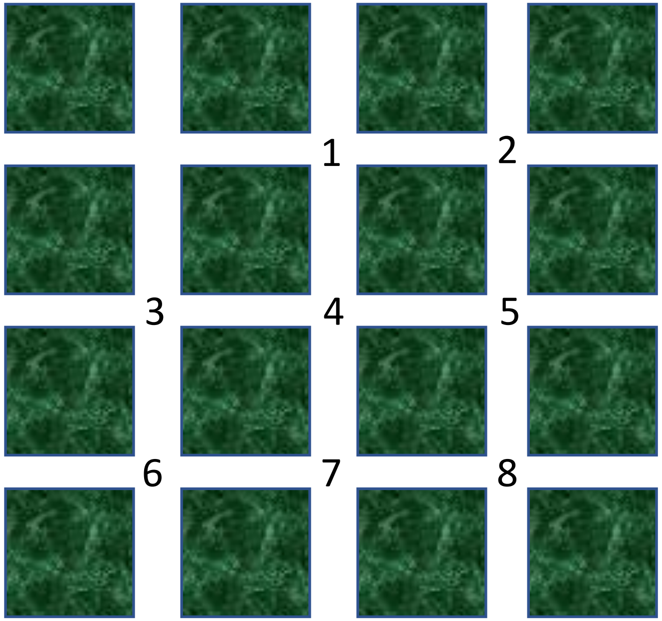
Probabilities of moving from old to new

Markov models: terminology

- *States*: intersections 1 – 8
- *Transition probability*: probability of moving from a given state to another state
- *Transition matrix*: any square matrix with non-negative entries where every column sums up to 1

	Old intersection							
New intersection	1	2	3	4	5	6	7	8
1	1/3	1/3	0	1/5	0	0	0	0
2	1/3	1/3	0	0	1/4	0	0	0
3	0	0	1/3	1/5	0	1/3	0	0
4	1/3	0	1/3	1/5	1/4	0	1/4	0
5	0	1/3	0	1/5	1/4	0	0	1/3
6	0	0	1/3	0	0	1/3	1/4	0
7	0	0	0	1/5	0	1/3	1/4	1/3
8	0	0	0	0	1/4	0	1/4	1/3

Probability vector for state 8



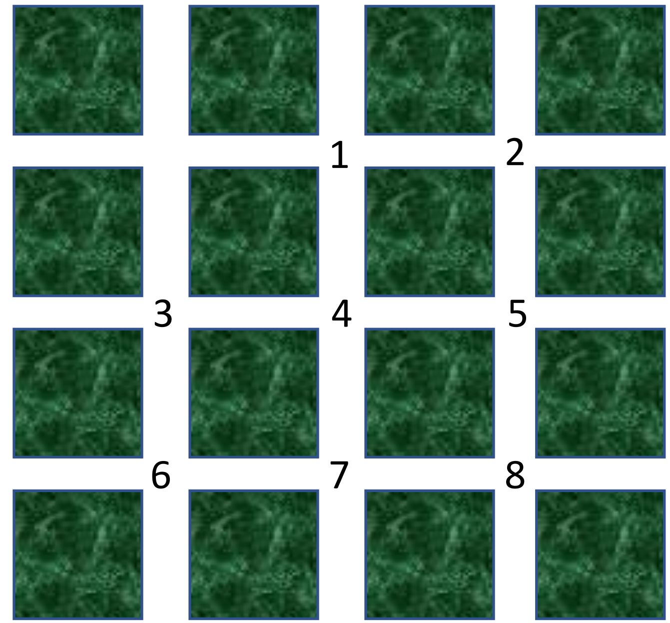
Markov model: theorem

If ***P*** is the transition matrix of a Markov process, and ***x****(t)* is a probability vector for state in time *t*, then the probability vector $\mathbf{x}(t+1) = \mathbf{P}^*\mathbf{x}(t)$

Intersection 5

We start with vector: **x**(0) = [0 0 0 0 0 1 0 0 0]

		Old intersection							
New intersection		1	2	3	4	5	6	7	8
	1	1/3	1/3	0	1/5	0	0	0	0
	2	1/3	1/3	0	0	1/4	0	0	0
	3	0	0	1/3	1/5	0	1/3	0	0
	4	1/3	0	1/3	1/5	1/4	0	1/4	0
	5	0	1/3	0	1/5	1/4	0	0	1/3
	6	0	0	1/3	0	0	1/3	1/4	0
	7	0	0	0	1/5	0	1/3	1/4	1/3
	8	0	0	0	0	1/4	0	1/4	1/3



$\mathbf{x}(1) = \mathbf{P}^*\mathbf{x}(0)$

$\mathbf{x}(2) = \mathbf{P}^*\mathbf{x}(1) = \mathbf{P}^{2*}x(0)$

$\mathbf{x}(3) = \mathbf{P}^*\mathbf{x}(2) = \mathbf{P}^{3*}\mathbf{x}(0)$

…

$\mathbf{x}(n) = \mathbf{P}^*\mathbf{x}(n\text{-}1) = \mathbf{P}^{n*}\mathbf{x}(0)$

The probabilities converge with time

- If the officer begins at intersection 5, his probable locations hour-by-hour are given in the table
- The probability vectors approach a fixed vector as *t* increases
- All the values after 22 hours will stay the same up to 3 decimal places

t	0	1	2	3	4	5	10	15	20	22
1	0	.000	.133	.116	.130	.123	.113	.109	.108	.107
2	0	.250	.146	.163	.140	.138	.115	.109	.108	.107
3	0	.000	.005	.039	.067	.073	.100	.106	.107	.107
4	0	.250	.113	.187	.162	.178	.178	.179	.179	.179
5	1	.250	.279	.190	.190	.168	.149	.144	.143	.143
6	0	.000	.000	.050	.056	.074	.099	.105	.107	.107
7	0	.000	.133	.104	.131	.125	.138	.141	.143	.143
8	0	.250	.146	.152	.124	.121	.108	.107	.107	.107

Check *markov.py*

Example: Migration

- A country is divided into 3 demographic regions.
- It is determined that each year:
  - Of the residents of region 1:
    - 5% move to region 2 and 5% move to region 3
  - Of the residents of region 2:
    - 15% move to region 1 and 10% move to region 3
  - Of the residents of region 3:
    - 10% move to region 1 and 5% move to region 2

- What percentage of the country resides in each region after a long period of time?

- Does it depend on the initial vector of proportion of residents in each region?

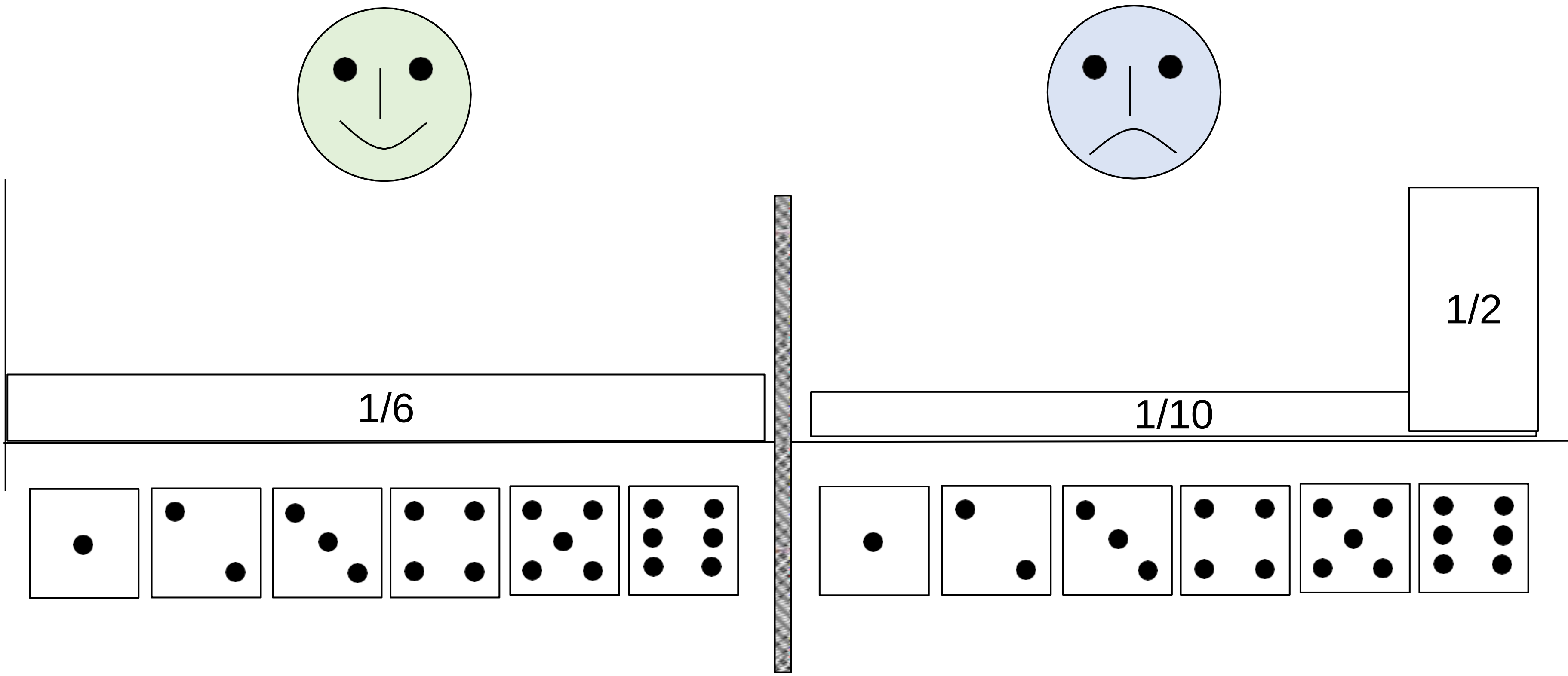
Markov models and sequences

Generating sequences from Markov models

See sample code in *casino.py*

The honest and the dishonest casino

Choose L with P(L) = 0.01



We assume that: P(F) = 0.99 P(L) = 0.01

Prior probabilities – before we see any evidence (sequence)

Bayes theorem and Markov models

- Pick a die at random - and roll
- We get 3 consecutive sixes: '666'
- Is the die loaded? What is the probability?

- We want to know P(L|3 sixes)
- From Bayes theorem:

P(L|3 sixes) = P(3 sixes|L)*P(L)/P(3 sixes)

P(F|3 sixes) = P(3 sixes|F)*P(F)/P(3 sixes)

The sequence was generated either by fair or by loaded die

P(3 sixes) = P(3 sixes|F)*P(F) + P(3 sixes|L) *P(L) = 0.0058

- P (**L**|3 sixes) = (0.5*0.5*0.5 * 0.01) /0.0058 = **0.215**
- P(**F**|3 sixes) = (1/6)*(1/6)*(1/6)*0.99 / 0.0058 = **0.785**

Not enough evidence to conclude that the die was Loaded

What are the odds?

- P (W1|evidence) = P(evidence|W1)*P(W1)/P(evidence)
- P (W2|evidence) = P(evidence|W2)*P(W2)/P(evidence)

- To compare P (W1|evidence) vs P (W2|evidence) :

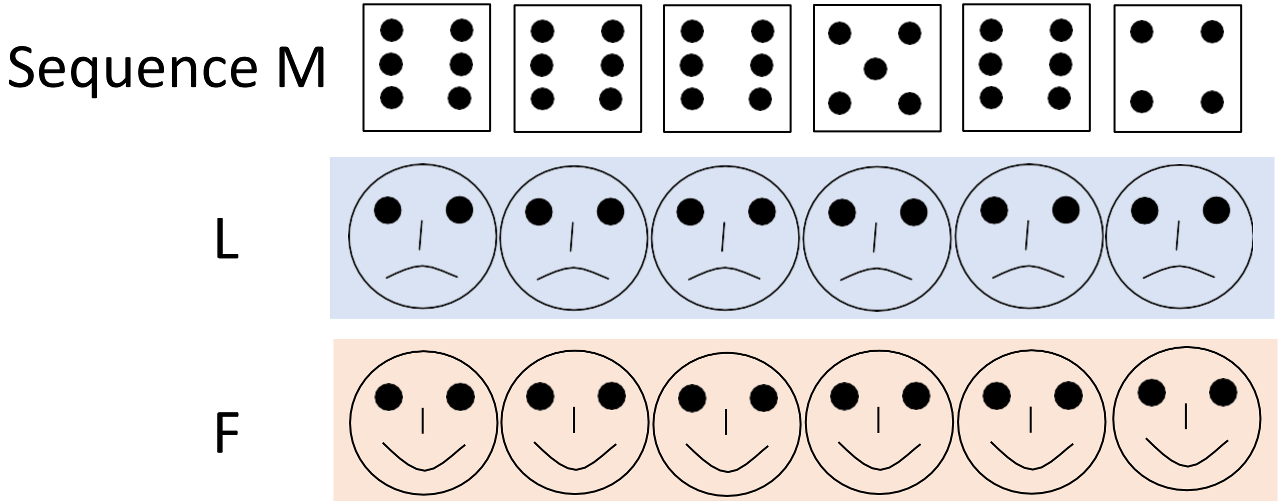
P (W1|evidence) / P (W2|evidence)

- Or to avoid underflow:

log [P (W1|evidence) / P (W2|evidence)]

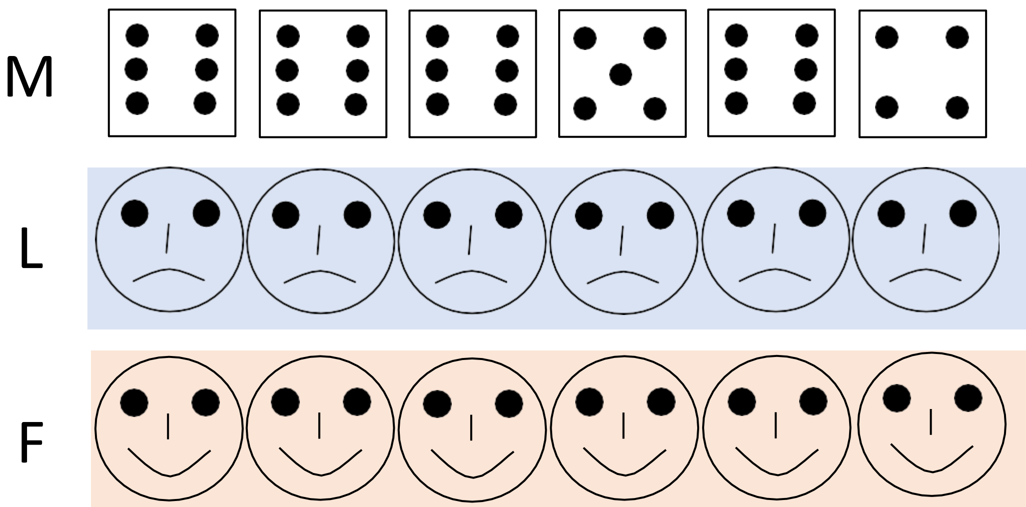
- Log odds ratio = log [P(evidence|W1)*P(W1)/ P(evidence|W2)*P(W2)]
- If > 0 – first is more likely, if < 0 – second is more likely

If two models **are <u>equally likely</u>**, we can use the conditional probabilities for discrimination



We can just compare P(M | L) and P(M | F)

We can use conditional probabilities for discrimination



OR

	F	L
1	0.17	0.10
2	0.17	0.10
3	0.17	0.10
4	0.17	0.10
5	0.17	0.10
6	0.17	0.50

P(M | L)=0.5*0.5*0.5*0.1*0.5*0.1=0.000625 = $6.25*10^{-4}$

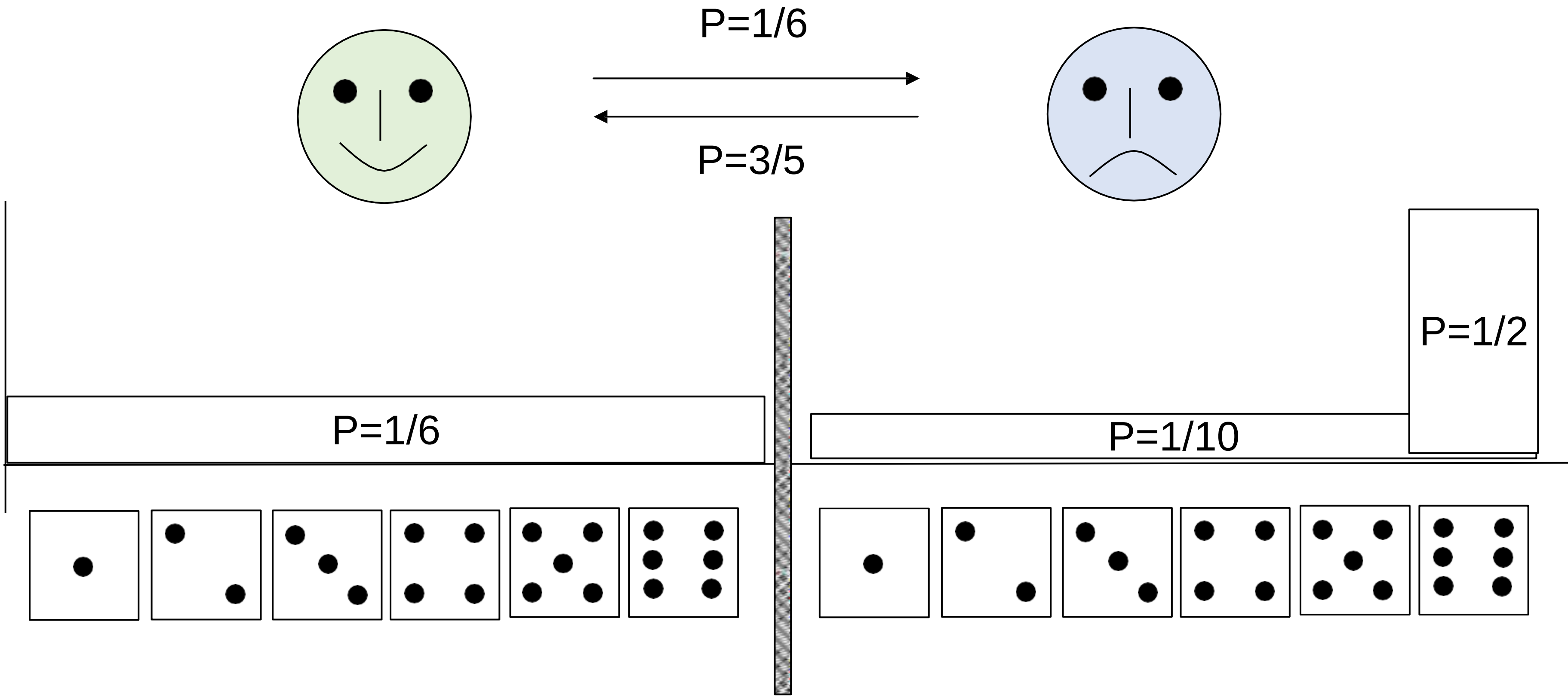
P(M | F)=0.17*0.17*0.17*0.17*0.17*0.17=0.000024 = $2.4 *10^{-5}$

How confident we are that this sequence was produced by a loaded die? P(M and model L)/ P(M and model F)=25.89

Or log [P(M I model L)/ P(M | F)]=1.4

Log-odds ratio

The occasionally dishonest casino



Sequence generated by a model of an occasionally dishonest casino