

21-317 MEETING 5: CONTINUOUS-TIME MARKOV CHAINS I EXPONENTIAL CLOCKS AND POISSONIZATION

MATTHEW ROSENZWEIG

ABSTRACT. Meeting 5 reviewed discrete-time Markov chains through transition and sample-path schematics, made the Markov property and the dual evolution of observables and laws explicit, and distinguished Markov memorylessness from independence. We then reviewed memoryless waiting-time distributions, explained what a rate measures, and ran a discrete chain on an independent Poisson clock. The meeting ended at the close of Section 8, after constructing $X_t = Y_{N_t}$, deriving the transition semigroup P_t , and showing how laws and stationary measures evolve. The infinitesimal generator and actual exit rates begin Meeting 6.

CONTENTS

6. Recap: the discrete-time process	2
6.1. Transition diagram, process, and sample paths	2
6.2. Stationarity does not imply detailed balance	2
6.3. The path law and Markov property	3
6.4. Processes, marginal laws, observables, and laws	3
7. Random clocks, memorylessness, and rates	4
7.1. Distribution reminders	4
7.2. Geometric and exponential memorylessness	4
7.3. What a rate measures	4
8. Poissonization	5
8.1. The Poisson clock	5
8.2. Running the chain on the clock	5
8.3. The transition semigroup	6
What to retain	6
Exercises	7
Sources and further reading	7
References	7

Date: September 2, 2026; revised September 4, 2026.

6. RECAP: THE DISCRETE-TIME PROCESS

6.1. Transition diagram, process, and sample paths. Let S be a finite state space. A discrete-time stochastic process is a random sequence

$$(6.1) \quad Y_0, Y_1, Y_2, \dots,$$

with values in S . One run produces one sample path

$$(6.2) \quad y_0, y_1, y_2, \dots$$

The process is random; the sample path is one realized sequence.

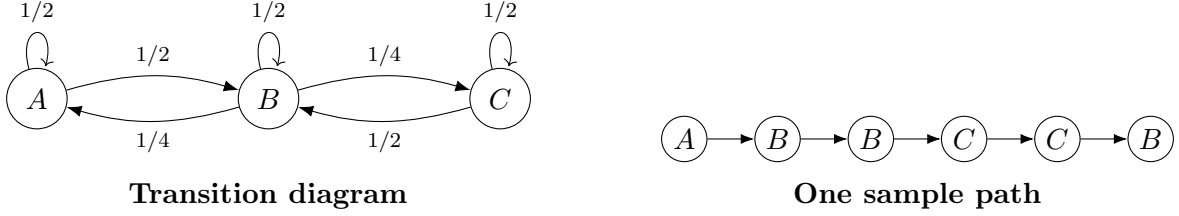


FIGURE 1. The transition diagram records all possible one-step moves and their probabilities. A sample path records the states actually visited in one realization.

With the state order (A, B, C) , the diagram corresponds to

$$(6.3) \quad K = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

Every row sums to one because the chain must select some next state.

Several possible realizations are

$$(6.4) \quad \begin{array}{c|ccccc} & n=0 & n=1 & n=2 & n=3 & n=4 \\ \hline \text{path 1} & A & B & B & C & C \\ \text{path 2} & C & C & B & A & A \\ \text{path 3} & B & A & B & C & C \end{array}.$$

Each column represents the random state at one time. Each row is one possible realized history. The transition diagram is the common random rule that generates such histories.

6.2. Stationarity does not imply detailed balance. On $S = \mathbb{Z}/3\mathbb{Z}$, consider the clockwise deterministic chain

$$(6.5) \quad K(i, i+1 \pmod{3}) = 1.$$

The uniform law $\mu(i) = 1/3$ is stationary:

$$(6.6) \quad (\mu K)(j) = \mu(j-1) = \frac{1}{3} = \mu(j).$$

It does not satisfy detailed balance because

$$(6.7) \quad \mu(0)K(0,1) = \frac{1}{3}, \quad \mu(1)K(1,0) = 0.$$

There is a persistent clockwise probability current. Stationarity balances total inflow and outflow at each state; detailed balance cancels every pair of opposite edge fluxes separately.

6.3. The path law and Markov property. Let ν_0 be an initial law and K a stochastic matrix on a finite state space S . The probability of a finite path is

$$(6.8) \quad \mathbb{P}(Y_0 = x_0, \dots, Y_n = x_n) = \nu_0(x_0) \prod_{j=0}^{n-1} K(x_j, x_{j+1}).$$

Proposition 6.1 (Markov property). *Whenever the conditioning event has positive probability,*

$$(6.9) \quad \mathbb{P}(Y_{n+1} = y \mid Y_0 = x_0, \dots, Y_n = x) = K(x, y).$$

More generally,

$$(6.10) \quad \mathbb{P}(Y_{n+m} = y \mid Y_0, \dots, Y_n) = K^m(Y_n, y).$$

Proof. Divide the path probability with terminal state y by the path probability through time n :

$$\begin{aligned} & \mathbb{P}(Y_{n+1} = y \mid Y_0 = x_0, \dots, Y_n = x) \\ &= \frac{\nu_0(x_0) \left(\prod_{j=0}^{n-1} K(x_j, x_{j+1}) \right) K(x, y)}{\nu_0(x_0) \prod_{j=0}^{n-1} K(x_j, x_{j+1})} = K(x, y). \end{aligned}$$

Summing over intermediate states gives (6.10). □

Two meanings of “memoryless”

Equation (6.9) says that, once the present state is known, the earlier history gives no further information about the future. It does not say that states at different times are independent. For example,

$$K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

defines a Markov chain with $Y_{n+1} = 1 - Y_n$.

6.4. Processes, marginal laws, observables, and laws. The process (Y_n) is a random sequence. Its marginal laws

$$(6.11) \quad \nu_n := \text{Law}(Y_n) = \nu_0 K^n$$

form a deterministic sequence. Marginal laws do not by themselves record temporal dependence; the joint path law (6.8) does.

For an observable $f : S \rightarrow \mathbb{R}$,

$$(6.12) \quad Kf(x) = \sum_y K(x, y) f(y) = \mathbb{E}[f(Y_{n+1}) \mid Y_n = x].$$

For a probability law ν ,

$$(6.13) \quad (\nu K)(y) = \sum_x \nu(x) K(x, y).$$

The actions are dual:

$$(6.14) \quad \langle \nu, Kf \rangle = \langle \nu K, f \rangle.$$

If $T : S \rightarrow S$ is deterministic and $K_T(x, y) = \mathbf{1}_{\{T(x)=y\}}$, then

$$(6.15) \quad K_T f = f \circ T = T^* f, \quad \nu K_T = T_\# \nu.$$

Thus a Markov kernel is a probabilistic generalization of a deterministic map.

7. RANDOM CLOCKS, MEMORYLESSNESS, AND RATES

7.1. Distribution reminders. Our convention is that a geometric random variable counts trials up to and including the first success.

Law	Values	Formula	Meaning
$B \sim \text{Bernoulli}(p)$	0, 1	$\mathbb{P}(B = 1) = p$	indicator of one event
$Z \sim \text{Bin}(n, p)$	$0, \dots, n$	$\mathbb{P}(Z = k) = \binom{n}{k} p^k (1-p)^{n-k}$	number of successes
$G \sim \text{Geom}(p)$	$1, 2, \dots$	$\mathbb{P}(G = k) = (1-p)^{k-1} p$	trials until first success
$T \sim \text{Exp}(a)$	$[0, \infty)$	$\mathbb{P}(T > t) = e^{-at}$	waiting time at rate a
$N \sim \text{Poisson}(\lambda)$	$0, 1, 2, \dots$	$\mathbb{P}(N = k) = e^{-\lambda} \lambda^k / k!$	count in an interval

Their means are

$$(7.1) \quad \mathbb{E}B = p, \quad \mathbb{E}Z = np, \quad \mathbb{E}G = \frac{1}{p}, \quad \mathbb{E}T = \frac{1}{a}, \quad \mathbb{E}N = \lambda.$$

For a Poisson variable, $\text{Var}(N) = \lambda$.

7.2. Geometric and exponential memorylessness. If $G \sim \text{Geom}(p)$, then

$$(7.2) \quad \mathbb{P}(G > m + n \mid G > m) = \mathbb{P}(G > n), \quad m, n \geq 0,$$

because $\mathbb{P}(G > k) = (1-p)^k$.

If $T \sim \text{Exp}(a)$, then

$$(7.3) \quad \mathbb{P}(T > s + t \mid T > s) = \mathbb{P}(T > t), \quad s, t \geq 0,$$

because the survival function is e^{-at} .

Conversely, if a continuous positive survival function $S(t)$ satisfies

$$(7.4) \quad S(s + t) = S(s)S(t),$$

then $g(t) = -\log S(t)$ is continuous and additive, so $g(t) = at$ and

$$(7.5) \quad S(t) = e^{-at}.$$

This explains why exponential waiting times are natural in a continuous-time Markov model: elapsed waiting time carries no additional information about the remaining wait.

7.3. What a rate measures. A rate has units

$$\frac{\text{events}}{\text{unit time}}.$$

The events depend on the model.

For a rate- a clock, the events are clock rings or update opportunities. If N_t counts rings by time t , then

$$(7.6) \quad \mathbb{E}N_t = at,$$

and the mean waiting time between rings is $1/a$. For a short time h ,

$$(7.7) \quad \mathbb{P}(\text{a ring during } [0, h]) = ah + O(h^2).$$

For a continuous-time chain currently at x , the transition rate $q(x, y)$ corresponds to actual jumps $x \rightarrow y$. The total exit rate

$$(7.8) \quad r(x) = \sum_{y \neq x} q(x, y)$$

corresponds to actual departures from x to any different state. A master clock may also ring without changing the state, so its ring rate need not equal the actual exit rate.

8. POISSONIZATION

8.1. The Poisson clock. Fix $a > 0$ and let

$$(8.1) \quad \tau_1, \tau_2, \dots$$

be independent exponential random variables of rate a . Define

$$(8.2) \quad T_0 = 0, \quad T_n = \tau_1 + \dots + \tau_n,$$

and

$$(8.3) \quad N_t = \max\{n : T_n \leq t\}.$$

Thus N_t counts clock rings by time t .

Proposition 8.1 (Poisson count law). *For $n = 0, 1, 2, \dots$,*

$$(8.4) \quad \mathbb{P}(N_t = n) = e^{-at} \frac{(at)^n}{n!}.$$

Consequently,

$$(8.5) \quad \mathbb{E}N_t = \text{Var}(N_t) = at.$$

Proof. For $n \geq 1$, the sum T_n has density

$$(8.6) \quad f_{T_n}(s) = \frac{a^n s^{n-1} e^{-as}}{(n-1)!}.$$

The event $\{N_t = n\}$ means that $T_n = s \leq t$ and the next wait exceeds $t - s$. Therefore

$$\begin{aligned} \mathbb{P}(N_t = n) &= \int_0^t f_{T_n}(s) e^{-a(t-s)} \, ds \\ &= e^{-at} \frac{a^n}{(n-1)!} \int_0^t s^{n-1} \, ds = e^{-at} \frac{(at)^n}{n!}. \end{aligned}$$

The case $n = 0$ is $\mathbb{P}(\tau_1 > t) = e^{-at}$. □

8.2. Running the chain on the clock. Let (Y_n) be a discrete-time Markov chain with transition kernel K , independent of the clock. Define

$$(8.7) \quad X_t := Y_{N_t}.$$

At every clock ring, one K -update occurs. Between rings, the state remains constant.

Poissonization replaces the deterministic update count n by the random number N_t of updates completed by time t .

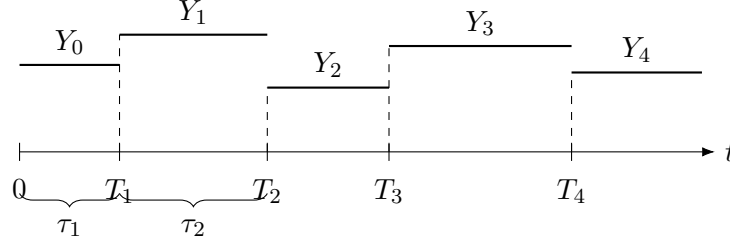


FIGURE 2. The exponential variables give the random gaps between update opportunities. The process $X_t = Y_{N_t}$ remains constant between clock rings.

8.3. The transition semigroup. For an observable f , define

$$(8.8) \quad P_t f(x) := \mathbb{E}_x[f(X_t)].$$

Conditioning on N_t gives

$$(8.9) \quad \begin{aligned} P_t f(x) &= \sum_{n=0}^{\infty} \mathbb{P}(N_t = n) K^n f(x) \\ &= e^{-at} \sum_{n=0}^{\infty} \frac{(at)^n}{n!} K^n f(x). \end{aligned}$$

This is a convex combination of stochastic matrices. Since I commutes with K ,

$$(8.10) \quad P_t = e^{-at} e^{atK} = e^{at(K-I)}.$$

Therefore

$$(8.11) \quad P_0 = I, \quad P_{s+t} = P_s P_t.$$

The family $(P_t)_{t \geq 0}$ is the transition semigroup of the continuous-time process.

If $\nu_0 = \text{Law}(X_0)$, then

$$(8.12) \quad \nu_t = \nu_0 P_t,$$

and

$$(8.13) \quad \langle \nu_0, P_t f \rangle = \langle \nu_t, f \rangle.$$

If $\mu K = \mu$, then $\mu P_t = \mu$ for every t .

WHAT TO RETAIN

- (1) The Markov property is conditional memorylessness of the future given the present; it does not mean independence across times.
- (2) The random process (Y_n) and its deterministic marginal laws $\nu_n = \nu_0 K^n$ are different objects.
- (3) The geometric and exponential distributions are the memoryless waiting-time laws in discrete and continuous time.
- (4) A rate counts a specified class of events per unit time. For the Poisson clock, the events are clock rings or update opportunities.
- (5) Poissonization defines

$$X_t = Y_{N_t}$$

and gives

$$P_t = e^{-at} \sum_{n \geq 0} \frac{(at)^n}{n!} K^n = e^{at(K-I)}.$$

(6) The transition operators satisfy

$$P_0 = I, \quad P_{s+t} = P_s P_t.$$

(7) If $\mu K = \mu$, then $\mu P_t = \mu$ for every $t \geq 0$.

EXERCISES

Exercise 8.2. Starting from the path law in Section 6, verify the Markov property and derive

$$\mathbb{P}(Y_{n+2} = z \mid Y_0, \dots, Y_n = x) = K^2(x, z).$$

Exercise 8.3. Prove the geometric and exponential memoryless identities from their survival functions.

Exercise 8.4. A clock rings at rate 3 per minute. Find the expected number of rings in two minutes, the mean waiting time, and the first-order probability of a ring during h minutes.

Exercise 8.5. Derive the Poisson-mixture formula for P_t by conditioning on N_t . Why is P_t stochastic?

Exercise 8.6. If $\mu K = \mu$, prove from the Poisson-mixture formula that $\mu P_t = \mu$ for every $t \geq 0$.

SOURCES AND FURTHER READING

For finite Markov chains, Poissonization, and continuous-time transition semigroups, see Levin and Peres [1] and Norris [2]. For the geometric, exponential, and Poisson distributions and the construction by exponential interarrival times, see Ross [3].

REFERENCES

- [1] D. A. Levin and Y. Peres, with E. L. Wilmer, *Markov Chains and Mixing Times*, 2nd ed., AMS, 2017, <https://pages.uoregon.edu/dlevin/MARKOV/mcmt2e.pdf>.
- [2] J. R. Norris, *Markov Chains*, Cambridge University Press, 1997.
- [3] S. M. Ross, *Introduction to Probability Models*, 12th ed., Academic Press, 2019.