Topic No. 8

1. **Markov Chain**

Expectation and Convergence

Jacob Bernoulli - a 17th century Swiss mathematician

Formalize a method of estimating the unknown probability of some event.

Bernoulli trial (or binomial trial)
Consider a random experiment with exactly two possible outcomes, **Success/Failure.**

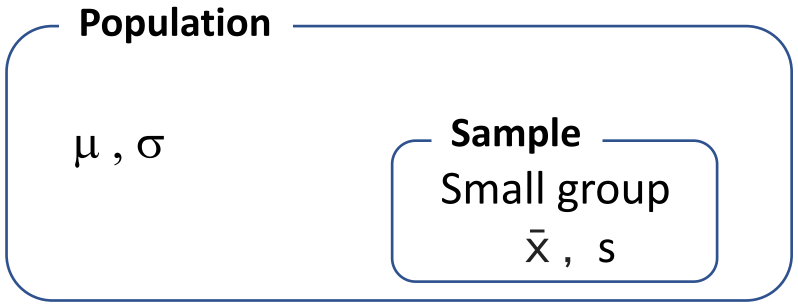
Examples (having only two possible outcomes)

Flipping a coin: Head denotes a success and Tail a failure. A fair coin has the probability of success 0.5.

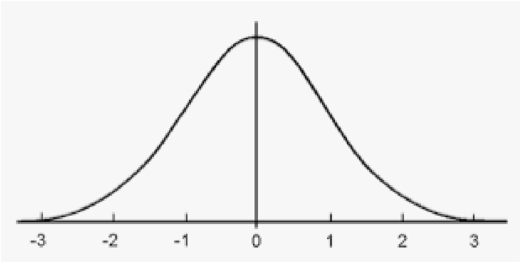
Rolling a die: A six denotes "success" and everything else a "failure". A fair dice has the probability of success 1/6.

Conducting a political opinion poll: A voter says he will vote "yes" or “no” in an upcoming referendum.

The Weak Law of Large Numbers:
Expected value **converges** on the actual ratio as the number of trials increases if conducted in an **independent** manner.



Central Limit Theorem:
Not only mean but variation follows a pattern (normal tail-like) – predetermined statistical fate



https://en.wikipedia.org/wiki/Bernoulli_trial

Francis Galton's Bean Machine
(Visit Boston Science Museum)

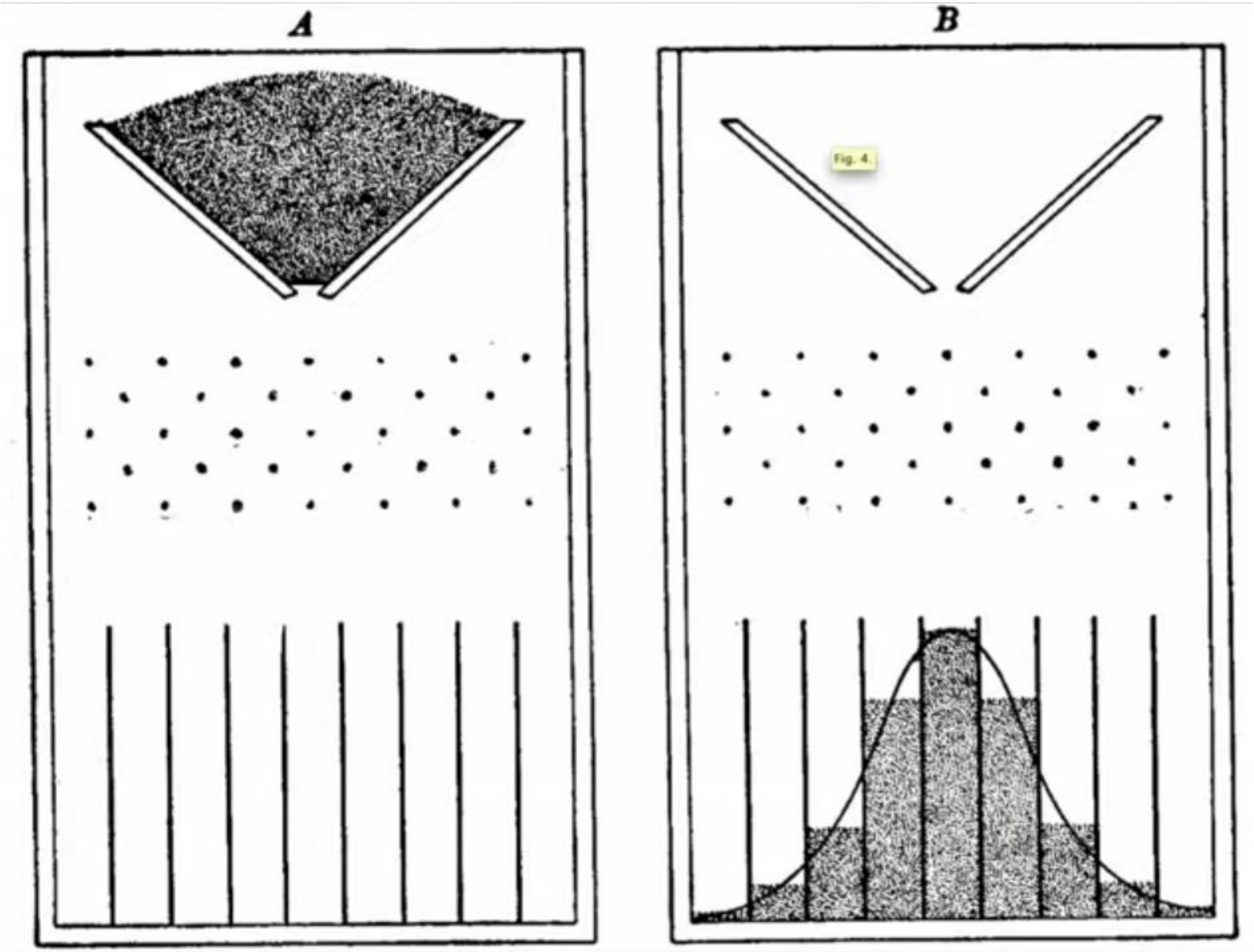
The number of times the event occurs in "independent" trials.

Refine the idea of Expectation – Can you quantify Expectation?

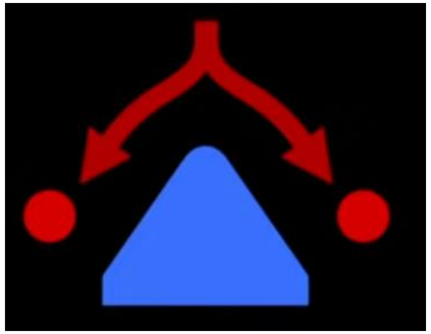
Pay attention to not only the expected values but also away from the expected value

Can the principles of independence be applied to forecasting tomorrow's weather?

Gender of two new born children in the neighborhood.



Each Collison → Independent Event



https://www.americanscientist.org/article/first-links-in-the-markov-chain

History behind Markov Chain

The Weak Law of Large Numbers:
Expected value converges on the actual ratio as the number of trials increases if done in independent manner.

?

Trials done in independent manner. → Expected value converges.

Jacob Bernoulli (1713)

Trials done in independent manner. → Expected value converges.

Pavel Nekrasov (1898)

"Independence is **a necessary condition** for the law of large numbers."

Religiously influenced mathematicians

"Events are inherently dependent and dependent events also converges following the law of large numbers."

Sun rises. → It gets warmer.

Trials done in **independent** manner. → Expected value converges

Trials done in **dependent** manner. → Expected value converges

Andrey Markov (1906)

Proving a weak law of large numbers without the independence assumption

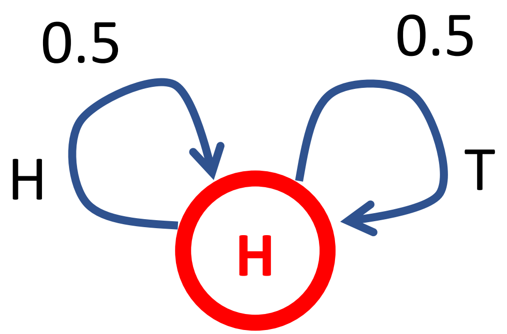
Independent vs. Dependent
(Do you see convergence in both cases?)

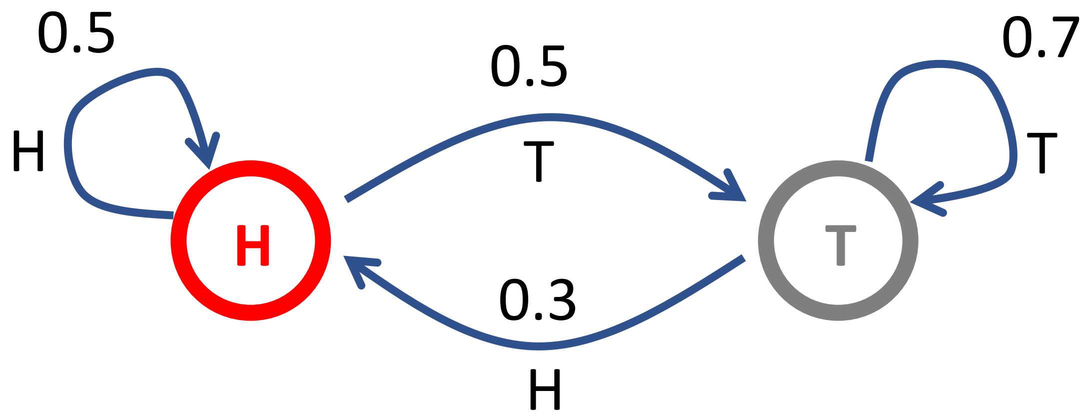
Property:
Probability of an event depends on occurrence of a previous event.

Trials done in **independent** manner.

Let the pattern size be 1000.
i.e., N = 1000

Trials done in **dependent** manner.





Flip a fair coin 1000 times.
Check the ratios:

- How many H’s over the total?
- How many T’s over the total?

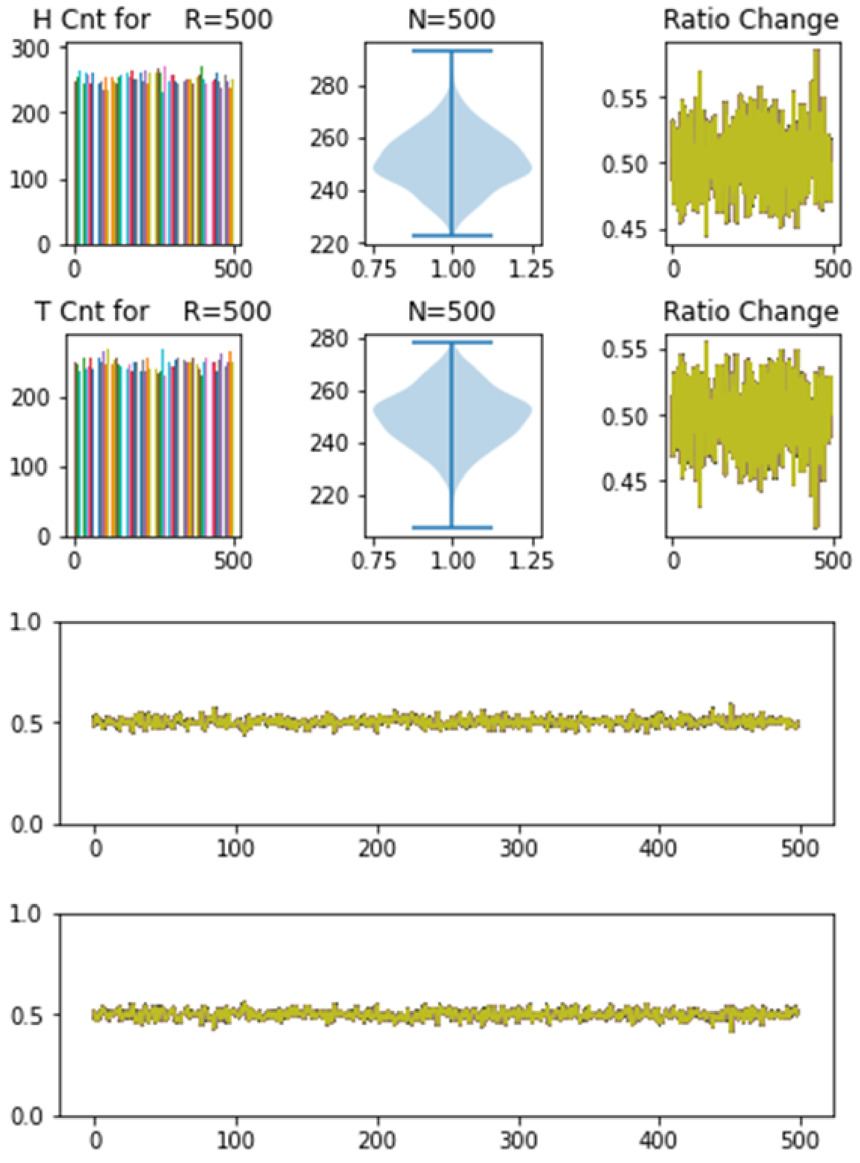
Two states are introduced where the outcome of flip dictates the state transition. Flip a fair coin 1000 times.
Check the ratios:

- How many H’s over the total?
- How many T’s over the total?

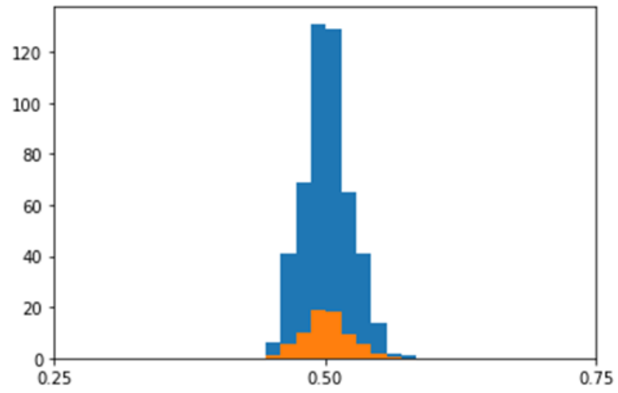
→ Same as how many times each state is visited.

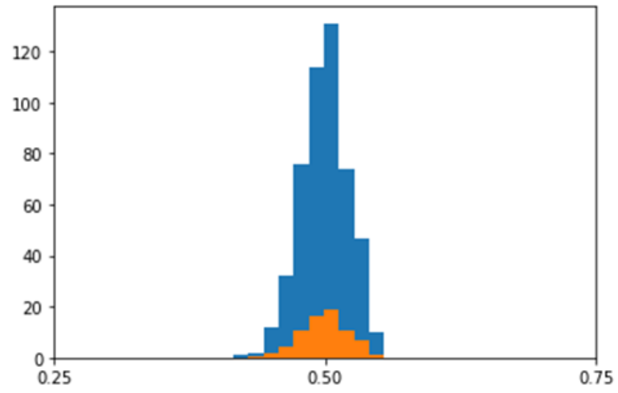
To be considered a proper Markov chain, a system must have a set of distinct states, with identifiable transitions between them.

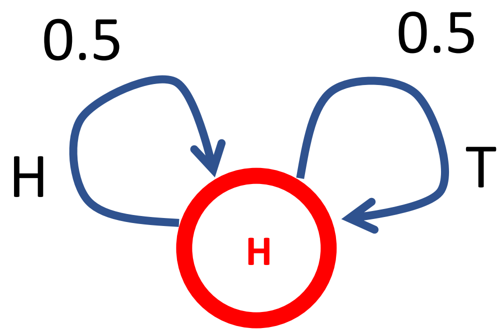
The Weak Law of Large Numbers / Central Limit Theorem (when events occur independently)



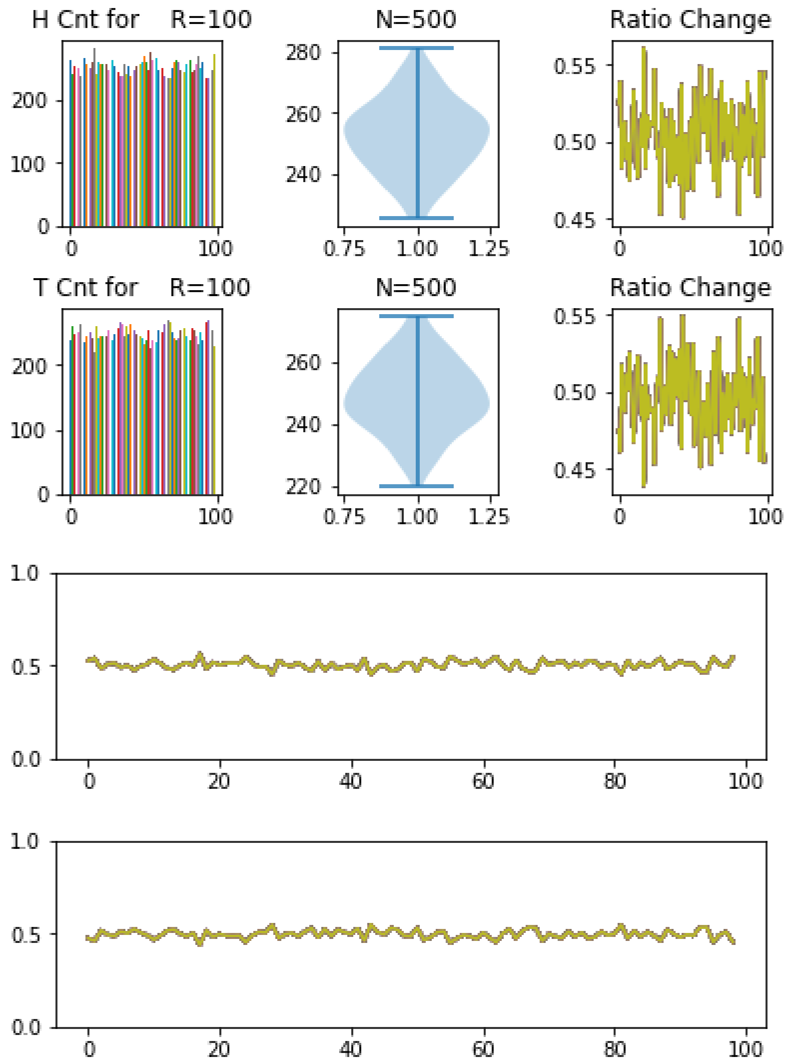
# Run the function with the data
R=500
N=500



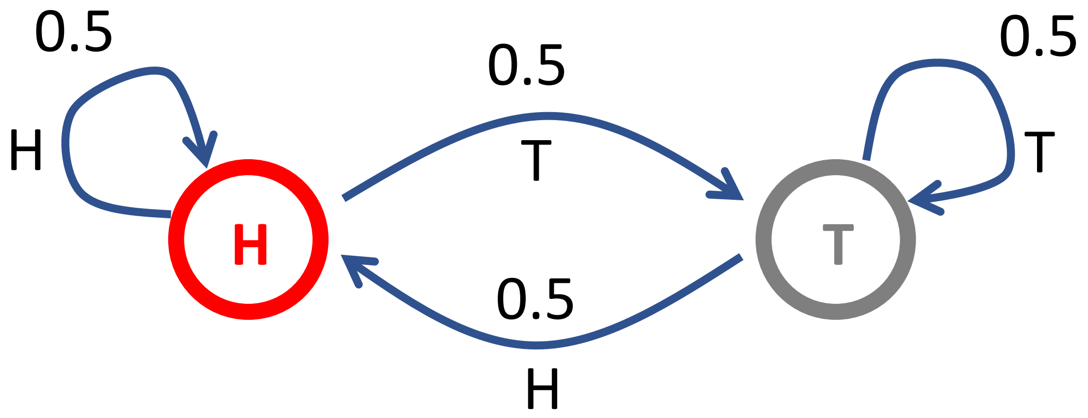




Markov Chain Stationary Distribution (when events occur dependently)



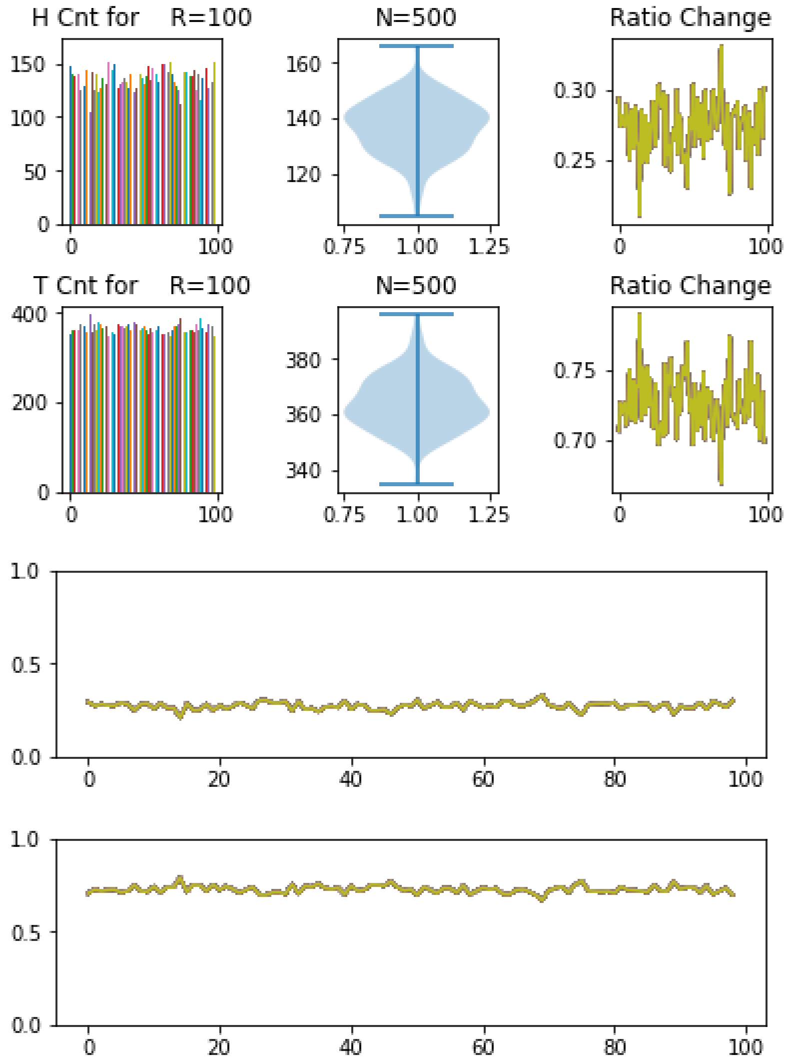
# Run the function with the data
R=100
N=500



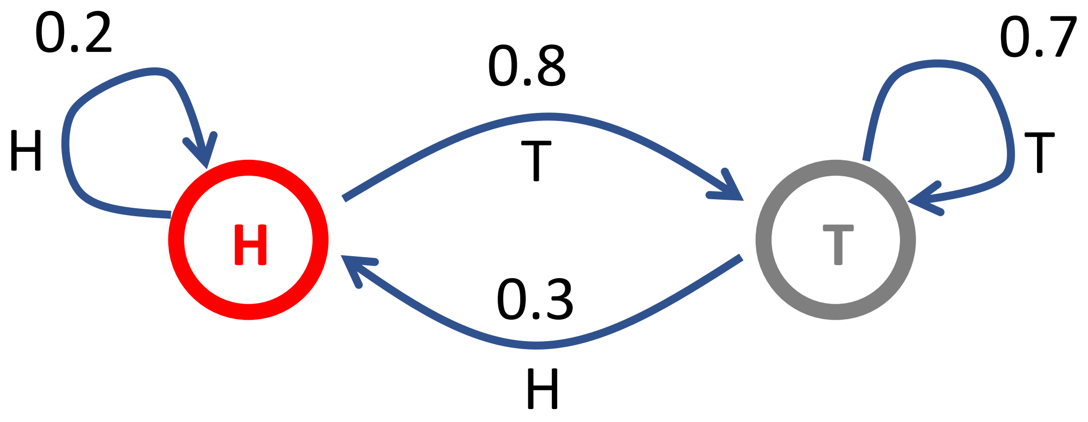
A **stationary distribution of a Markov chain** is a probability distribution that remains unchanged in the Markov chain as time progresses.

https://brilliant.org/wiki/stationary-distributions/

Markov Chain Stationary Distribution (when events occur dependently)



# Run the function with the data
R=100
N=500

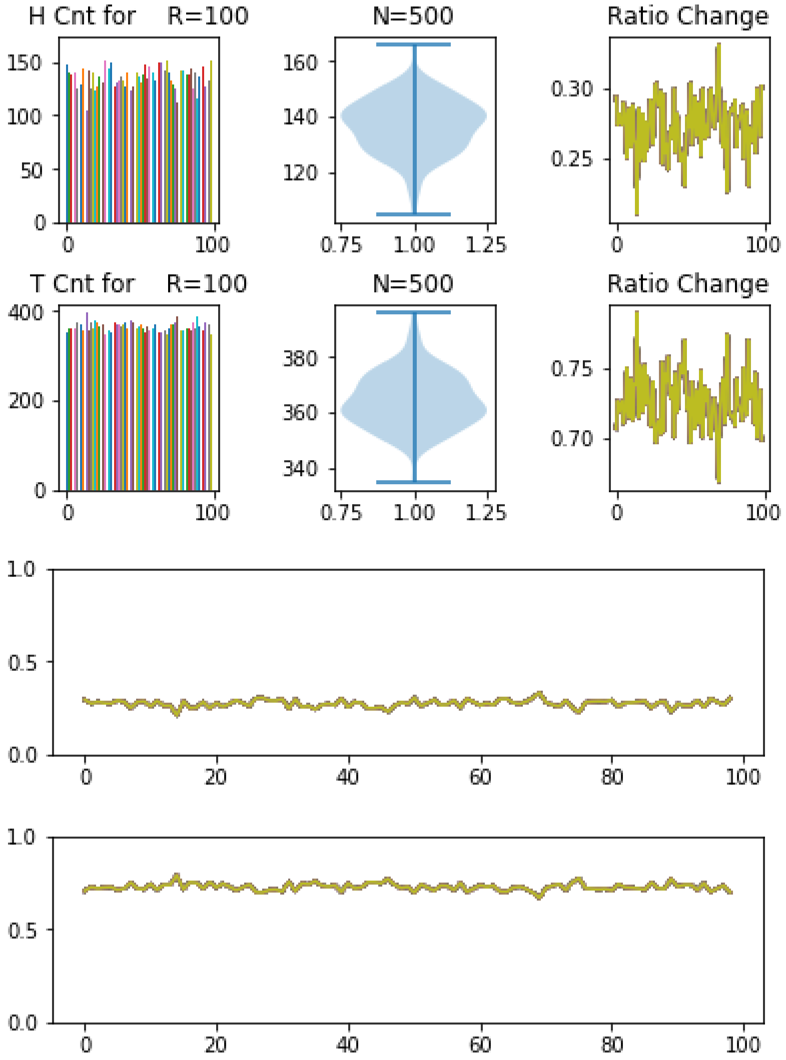


A **stationary distribution of a Markov chain** is a probability distribution that remains unchanged in the Markov chain as time progresses.

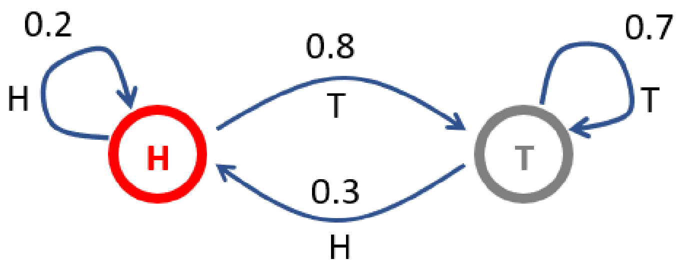
https://brilliant.org/wiki/stationary-distributions/

Markov Chain Stationary Distribution (when events occur dependently)

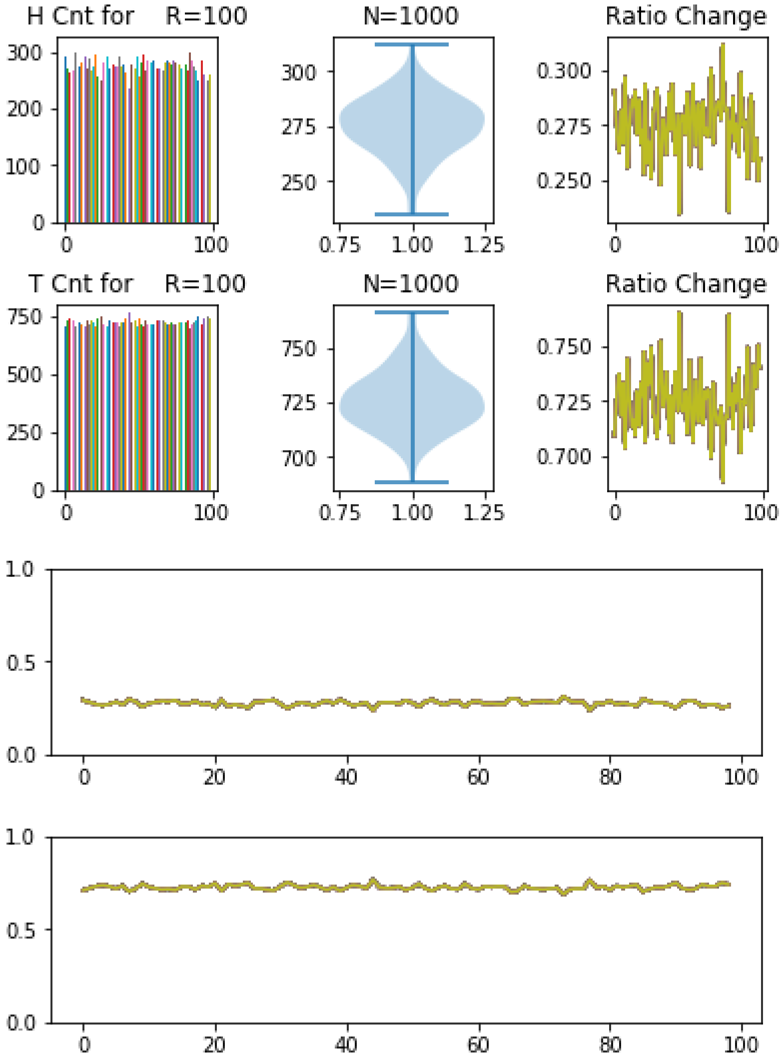
140 vs. 360



R=100
N=500



280 vs. 720



R=100
N=1000

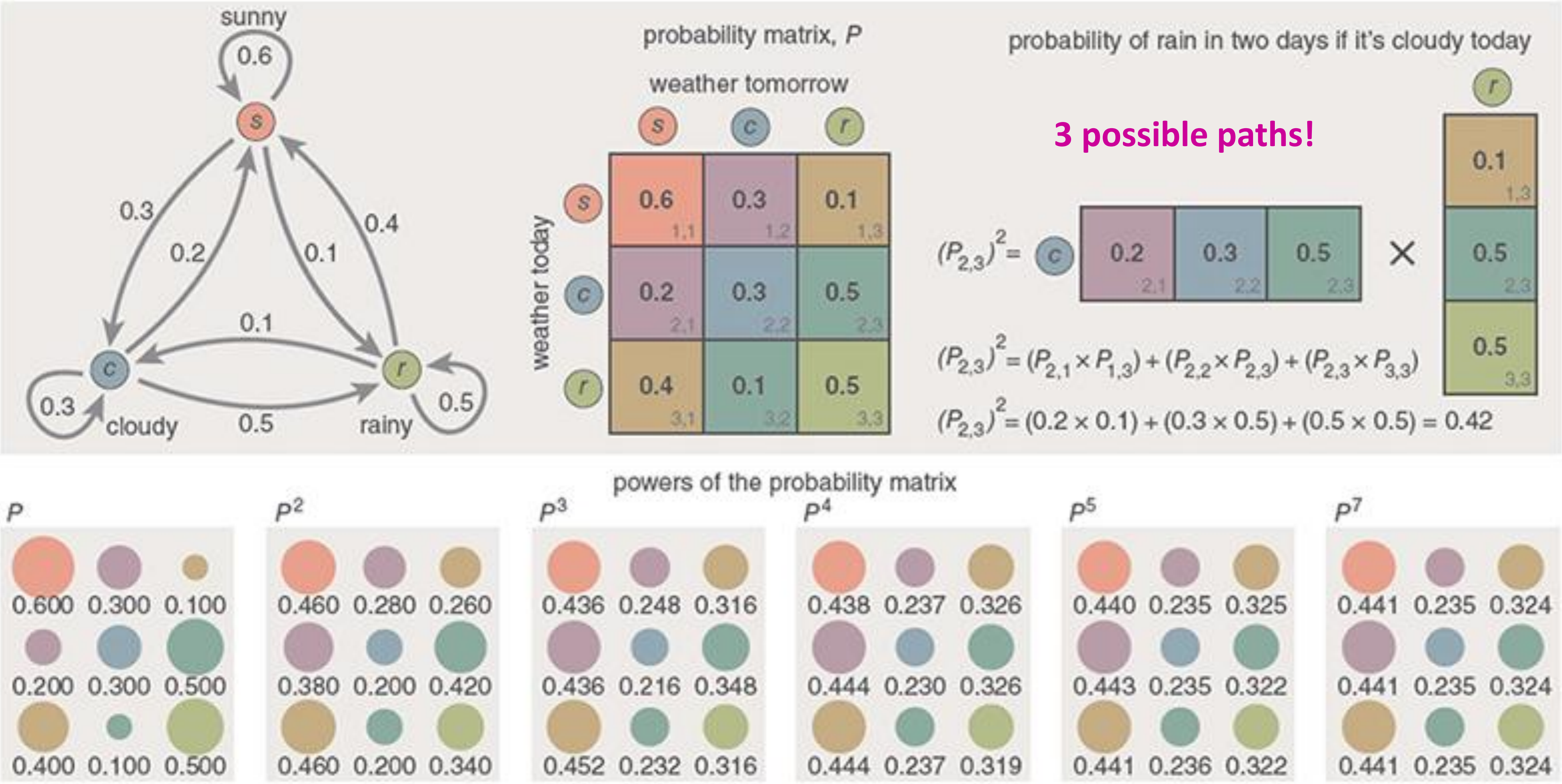
https://brilliant.org/wiki/stationary-distri

A **stationary distribution of a Markov chain**

Markov Chain

The chain can also answer questions such as, **“If it’s cloudy today, what is the probability of rain two days from now?”** The answer is found by summing the contributions of all pathways that lead from the cloudy state to the rainy state in exactly two steps. This sounds like a tedious exercise, but there’s an easy way to organize the computation, based on the arithmetic of matrices.

To calculate the probability of this specific sequence, just multiply the probabilities associated with the corresponding transition arrows.



https://www.americanscientist.org/article/first-links-in-the-markov-chain

Illustration by Brian Hayes.

Limitations to Markov Model

If the time interval is too short, then Markov models are inappropriate because the individual displacements are not random, but rather are **deterministically related in time**. This example suggests that Markov models are generally inappropriate over sufficiently short time intervals.

Timothy DelSole, “A Fundamental Limitation of Markov Models”, July 1, 2000, Journal of the Atmospheric Sciences, Volume 57: Issue 13

Boss who has been appointed for political reason. vs. Boss who has been promoted through the ranks.