

Likelihood and Bayesian inference

Michael Matschiner

Ludwig Maximilian University of Munich

Inference methods

1. Maximum likelihood

2. Bayesian inference

**Georgia Tsambos,
Martin Petr**



3. Simulation-based inference

4. Machine learning

Andrew Kern



Likelihood

Probability

Probability



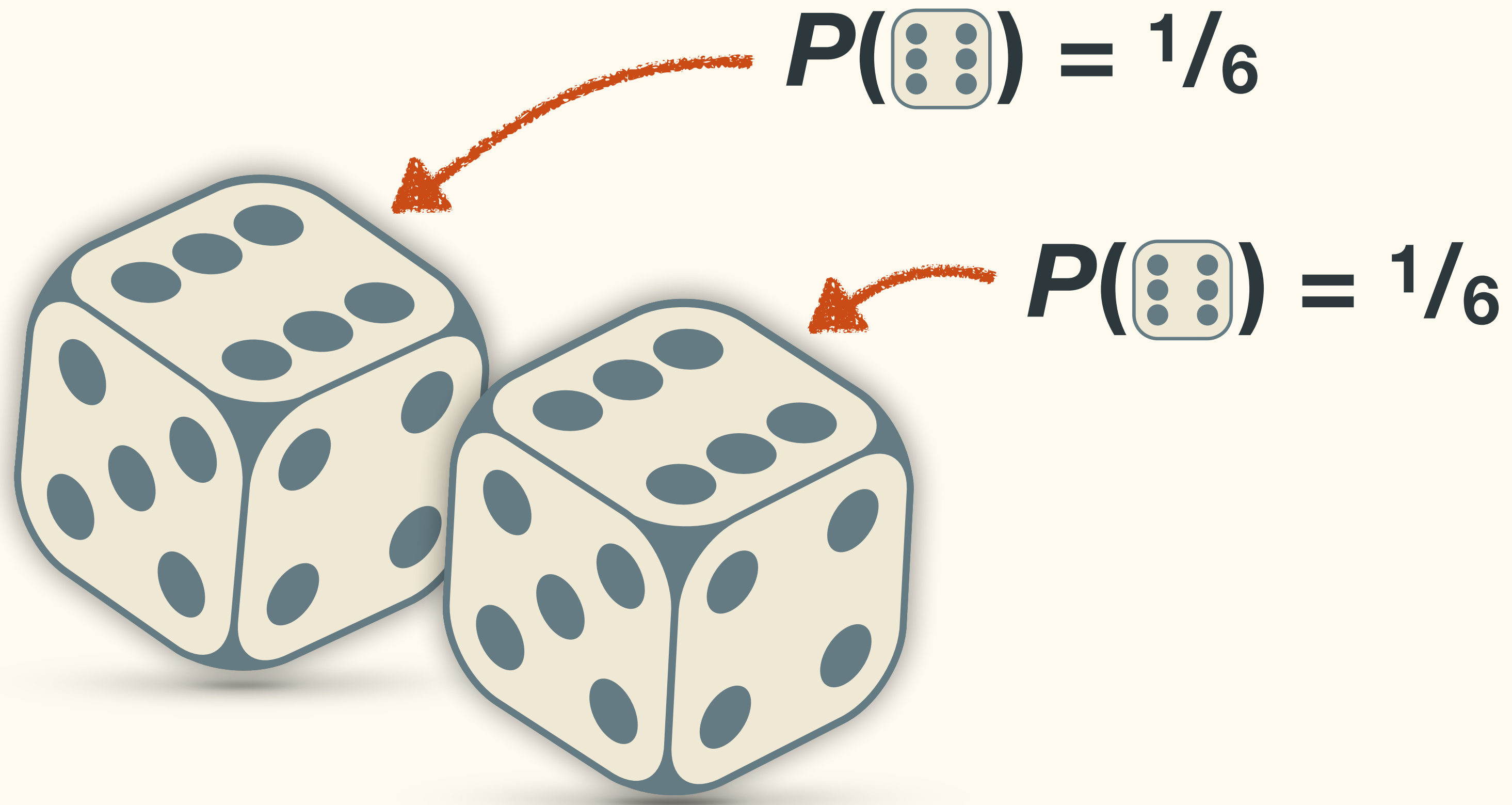
$$P = 1/6$$

Probability



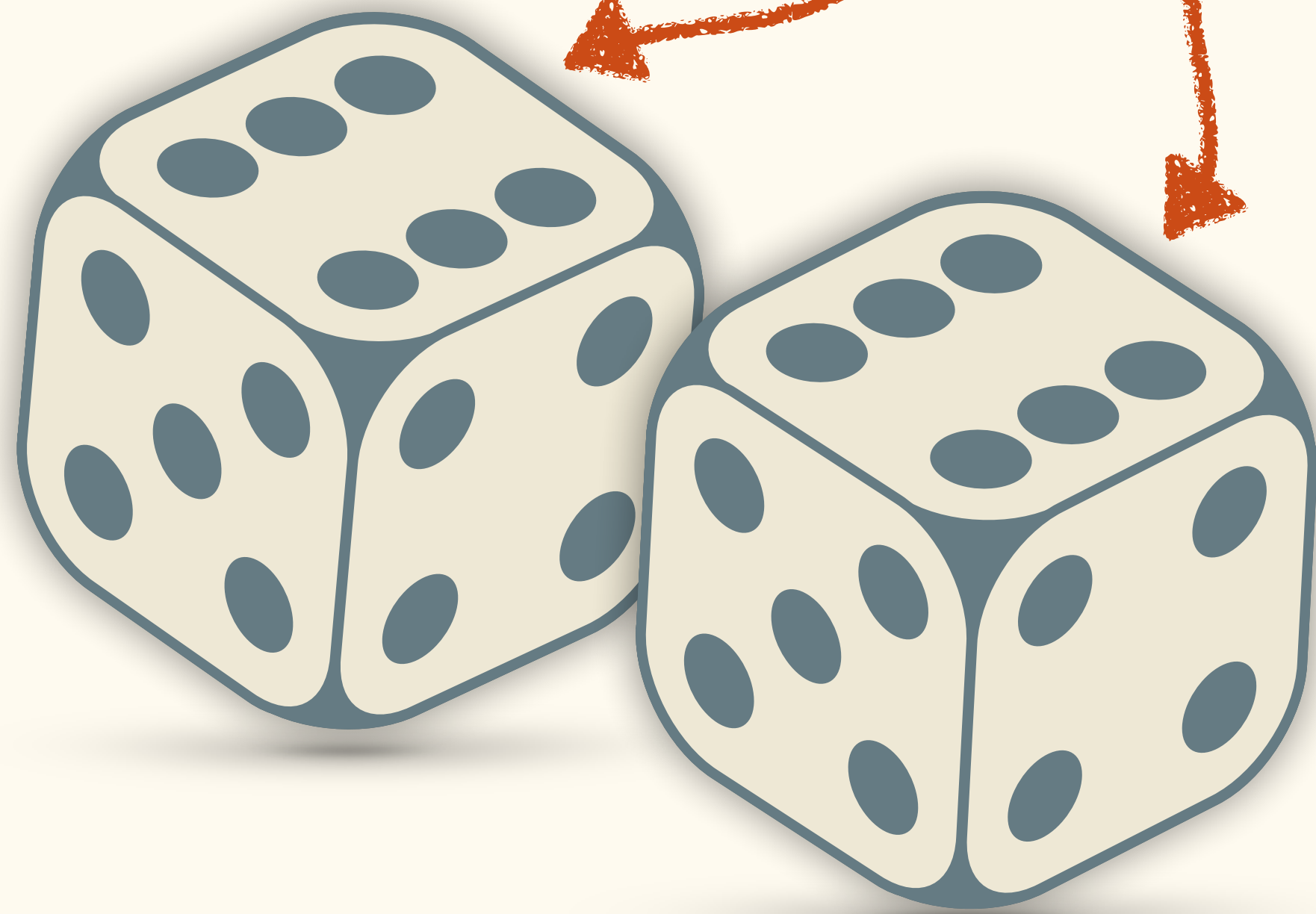
$$P(\text{⊞}) = 1/6$$

Probability



Probability

$$P(\text{Ⓜ} \& \text{Ⓜ}) = 1/6 \times 1/6 = 1/36$$



Probability



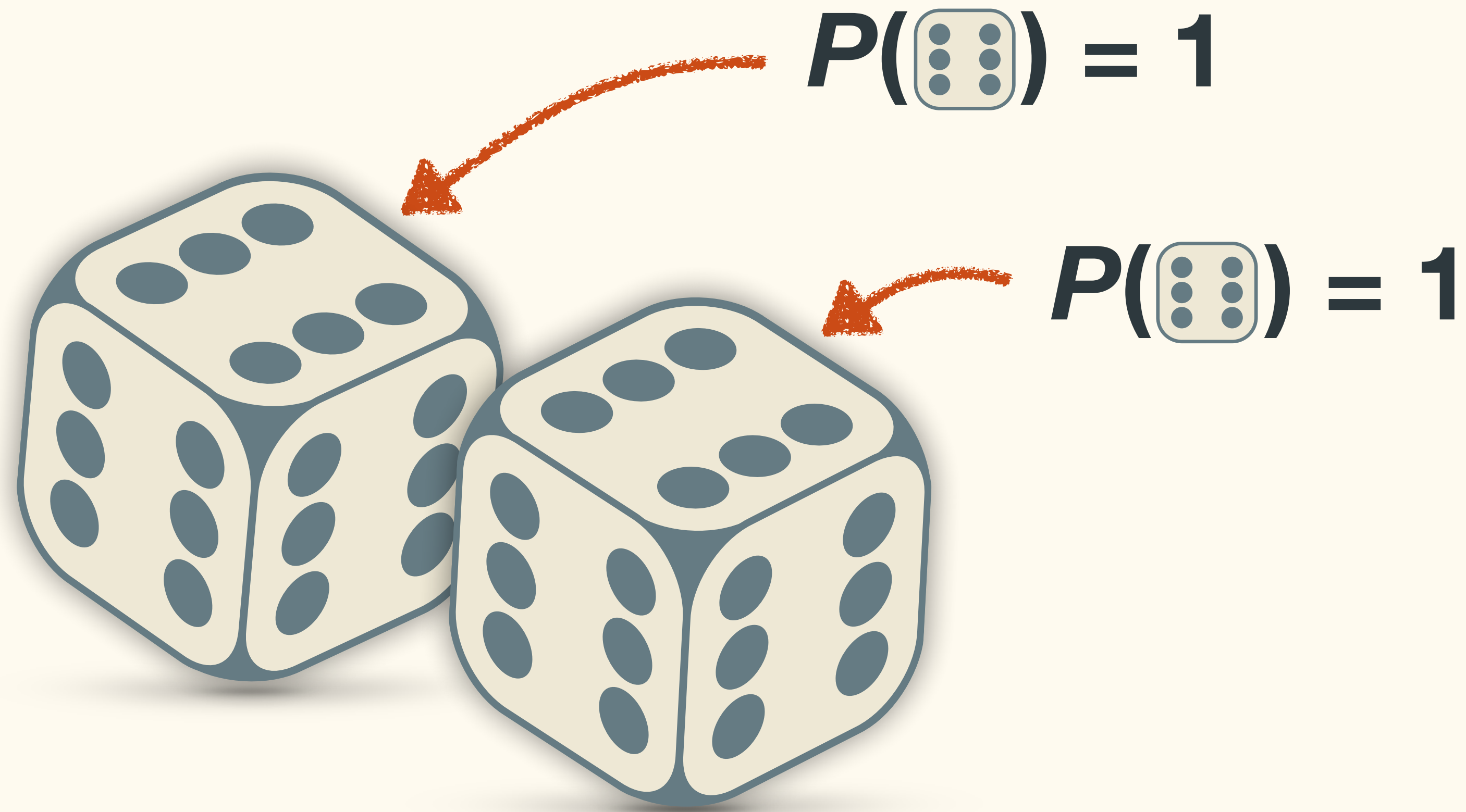
$$P(\text{Ⓜ}) = ?$$

Probability



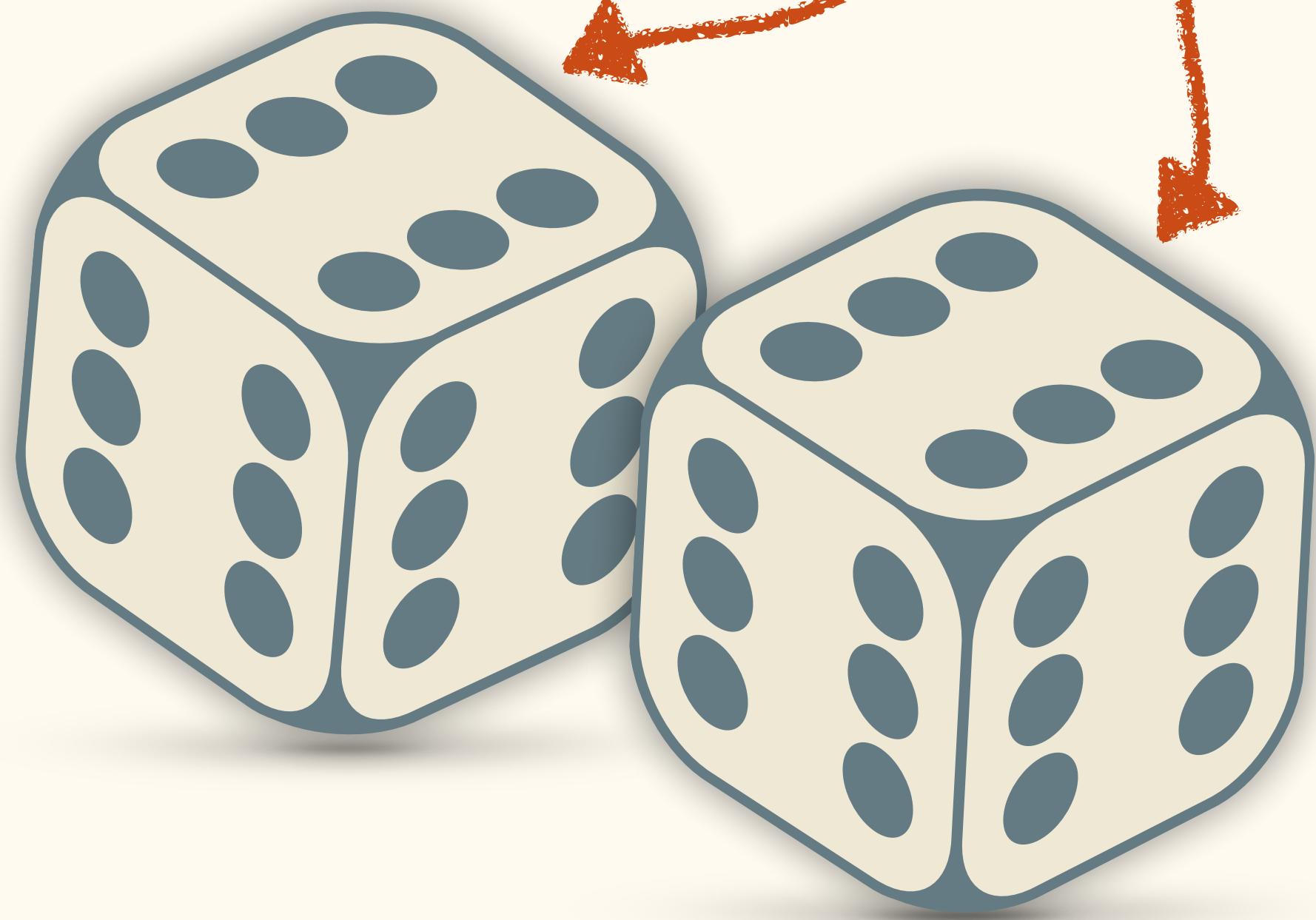
$$P(\text{Ⓜ}) = 1$$

Probability



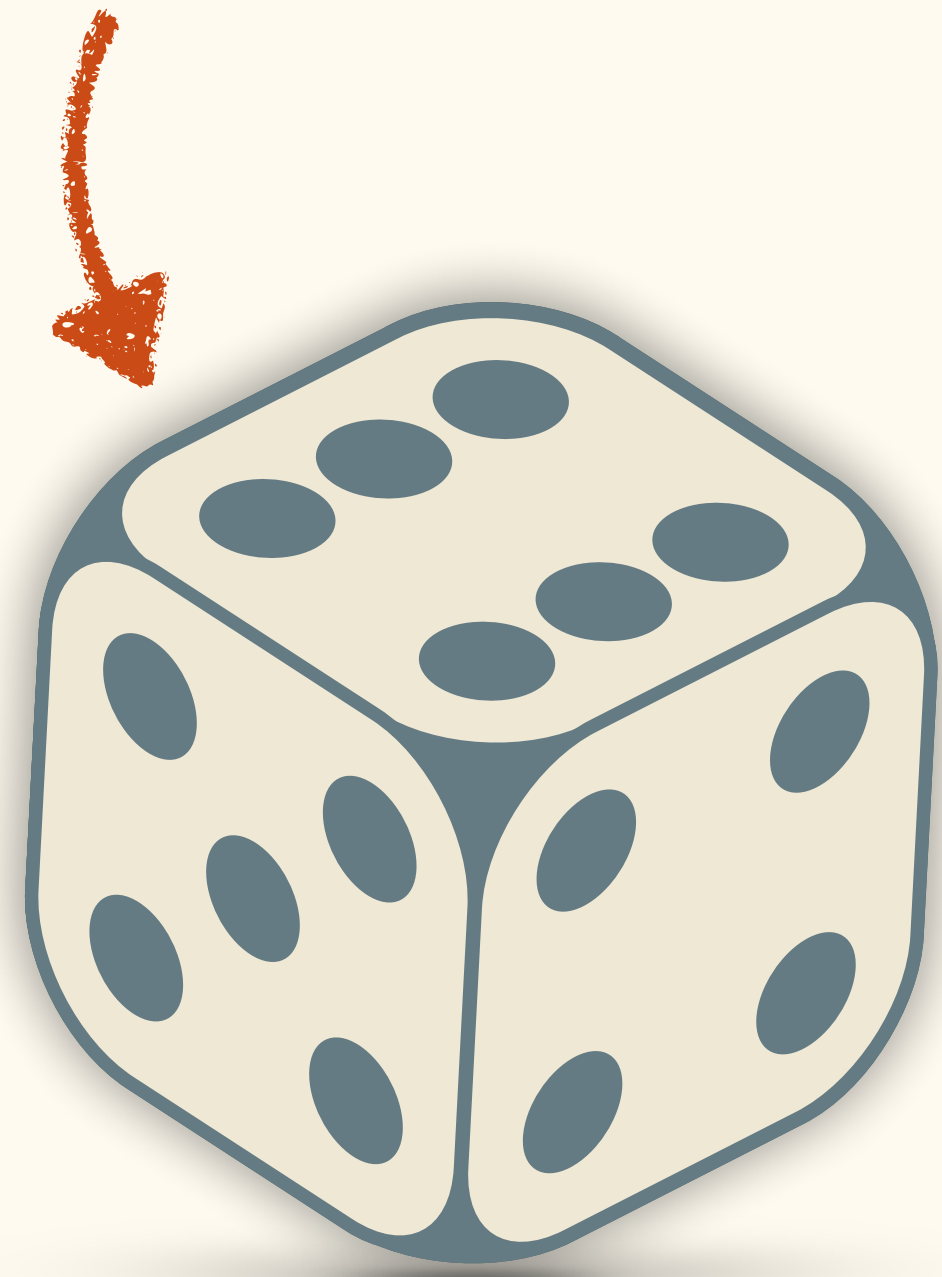
Probability

$$P(\text{Ⓛ} \& \text{Ⓛ}) = 1 \times 1 = 1$$

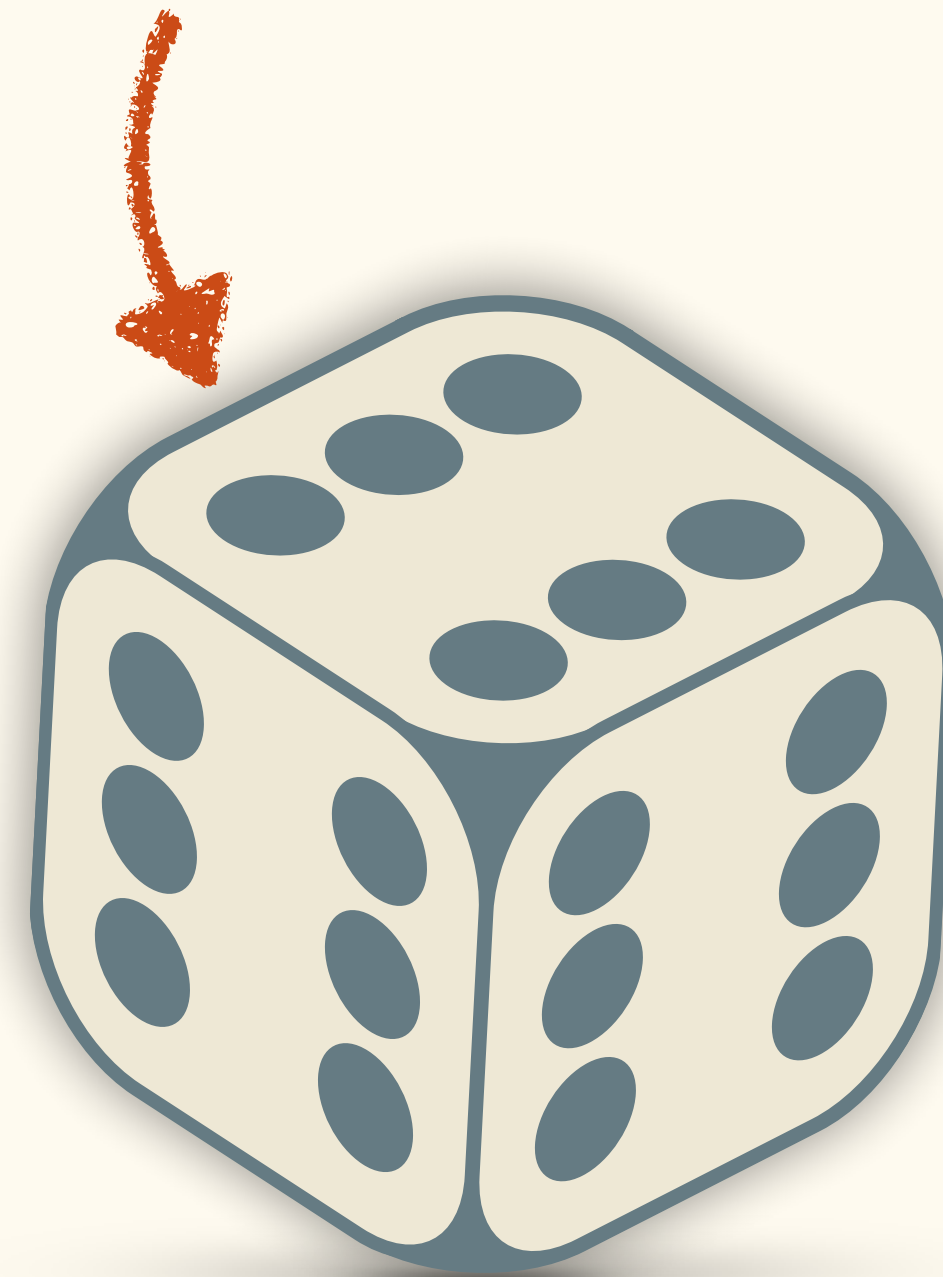


Probability

$$P(\text{⊠}) = 1/6$$



$$P(\text{⊠}) = 1$$



Probability

$$P(\text{🎲} \mid \text{🎲}) = 1/6$$



Observation

Probability

$$P(\text{🎲} \mid \text{🎲}) = 1/6$$



Result

Probability

“given” **“under the assumption of”**

$$P(\text{Result} \mid \text{Model}) = 1/6$$

Result **Model**
(fair dice)

Probability

Fair dice

$$P(\text{3 dots} \mid \text{Fair dice}) = 1/6$$

$$P(\text{3 dots} \mid \text{Trick dice}) = 1$$

Trick dice

Likelihood

Fair dice

$$L(\text{Fair dice} \mid \text{3, 3, 3}) = 1/6$$

Trick dice

$$L(\text{Trick dice} \mid \text{3, 3, 3}) = 1$$

Likelihood

$$L(\text{die} \mid \text{roll}) = P(\text{roll} \mid \text{die})$$

Probability



$P(A) = ?$

Probability

$$P(\mathbf{A}) = 1/4$$



$$P(\mathbf{A}) = 1$$



Probability

A, C, G, T

$$P(A \mid \triangle_{AC}) = 1/4$$


$$P(A \mid \triangle_{AA}) = 1$$

Only A



Likelihood

A, C, G, T


$$L(\text{die with A and C} \mid A) = 1/4$$

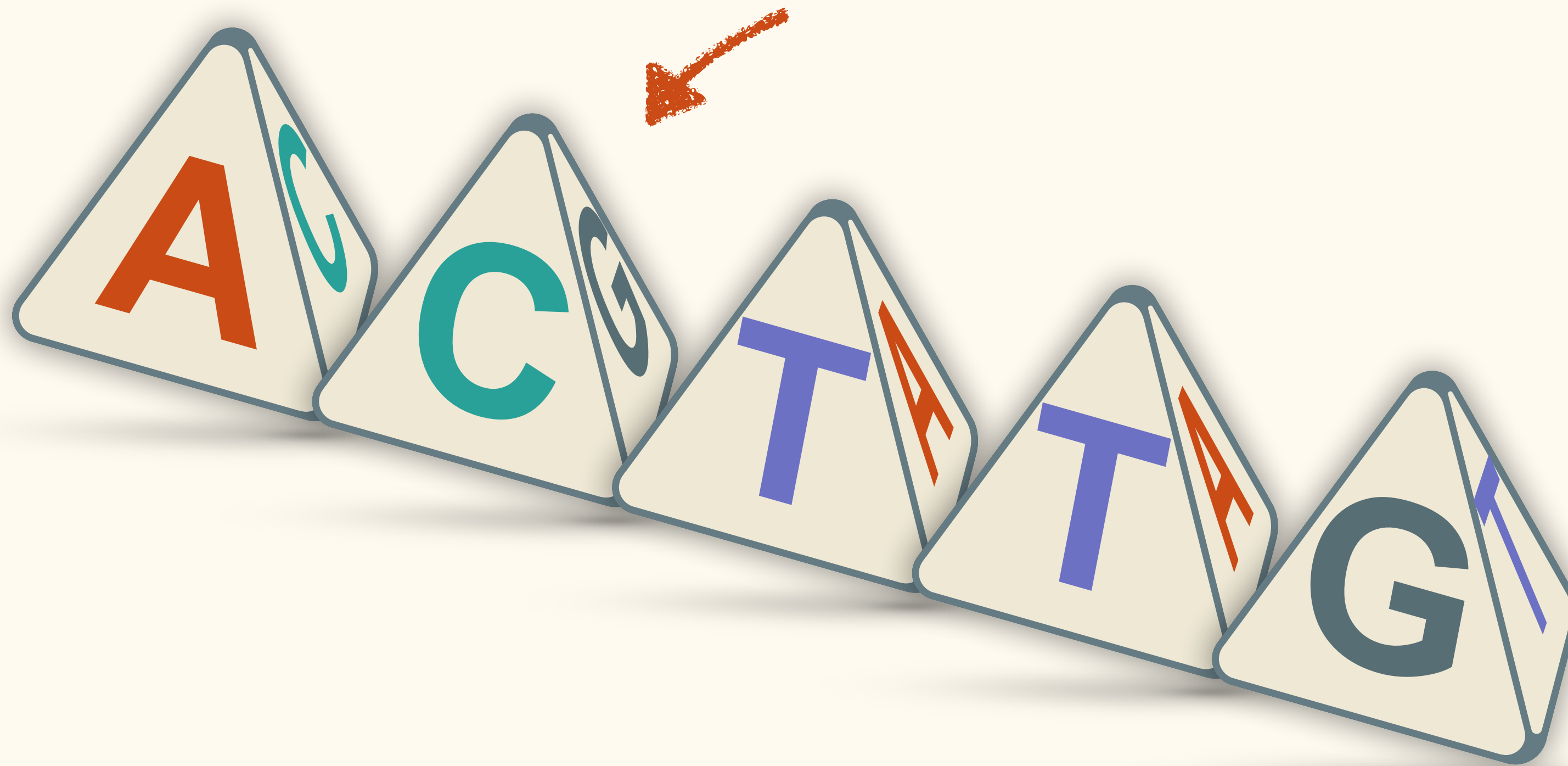
$$L(\text{die with A and A} \mid A) = 1$$

Only A



Probability

$$P(\text{ACTTG}) = ?$$



Probability

A, C, G, T

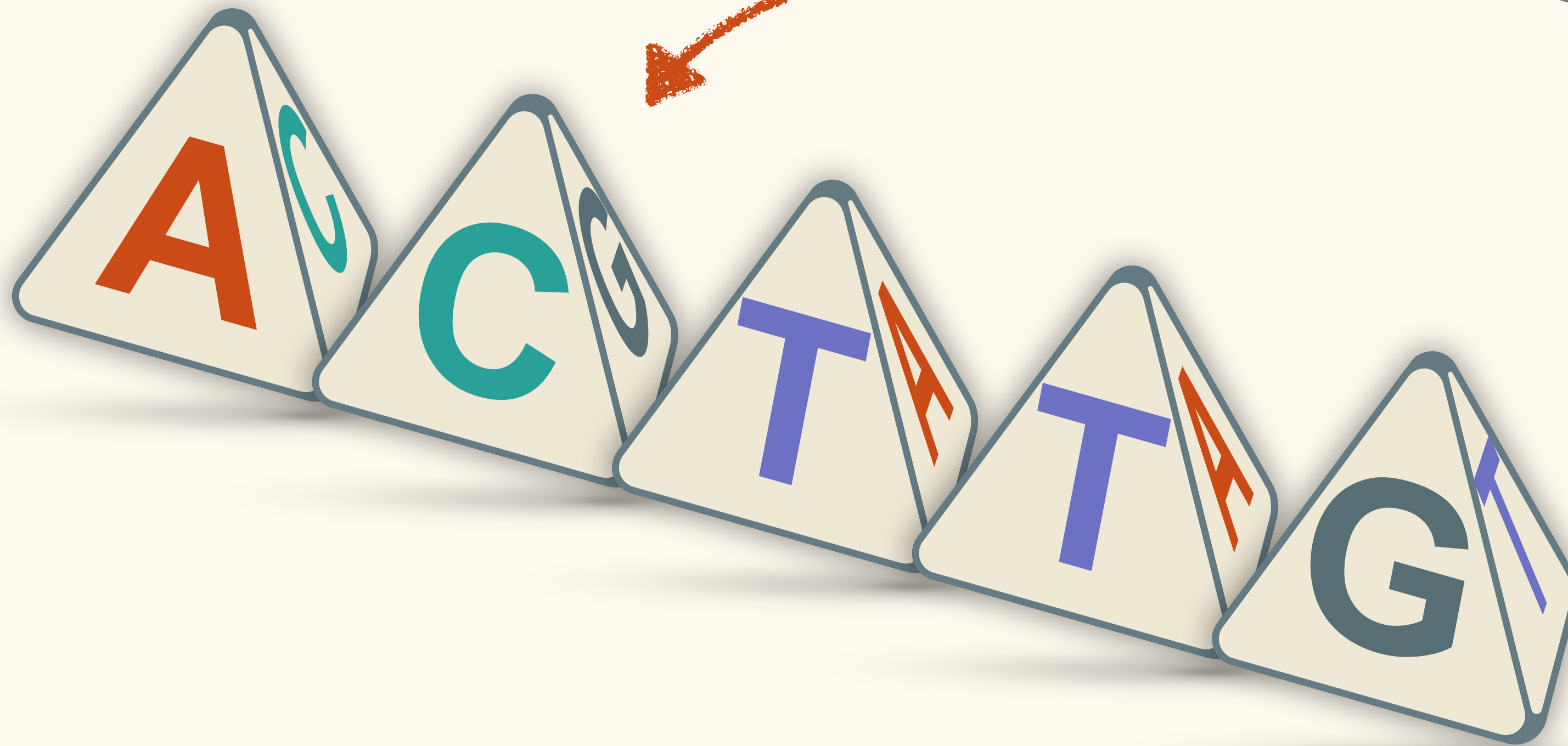
$$P(\text{ACTTG} \mid \triangle_{AC}) = ?$$

The image illustrates a probability problem using dice. Five dice are arranged in a line, showing the sequence ACTTG. The first die shows A and C, the second shows C and G, the third shows T and A, the fourth shows T and A, and the fifth shows G and T. An arrow points from the text 'A, C, G, T' to the first die, and another arrow points from the text 'P(ACTTG | triangle_{AC}) = ?' to the first die.

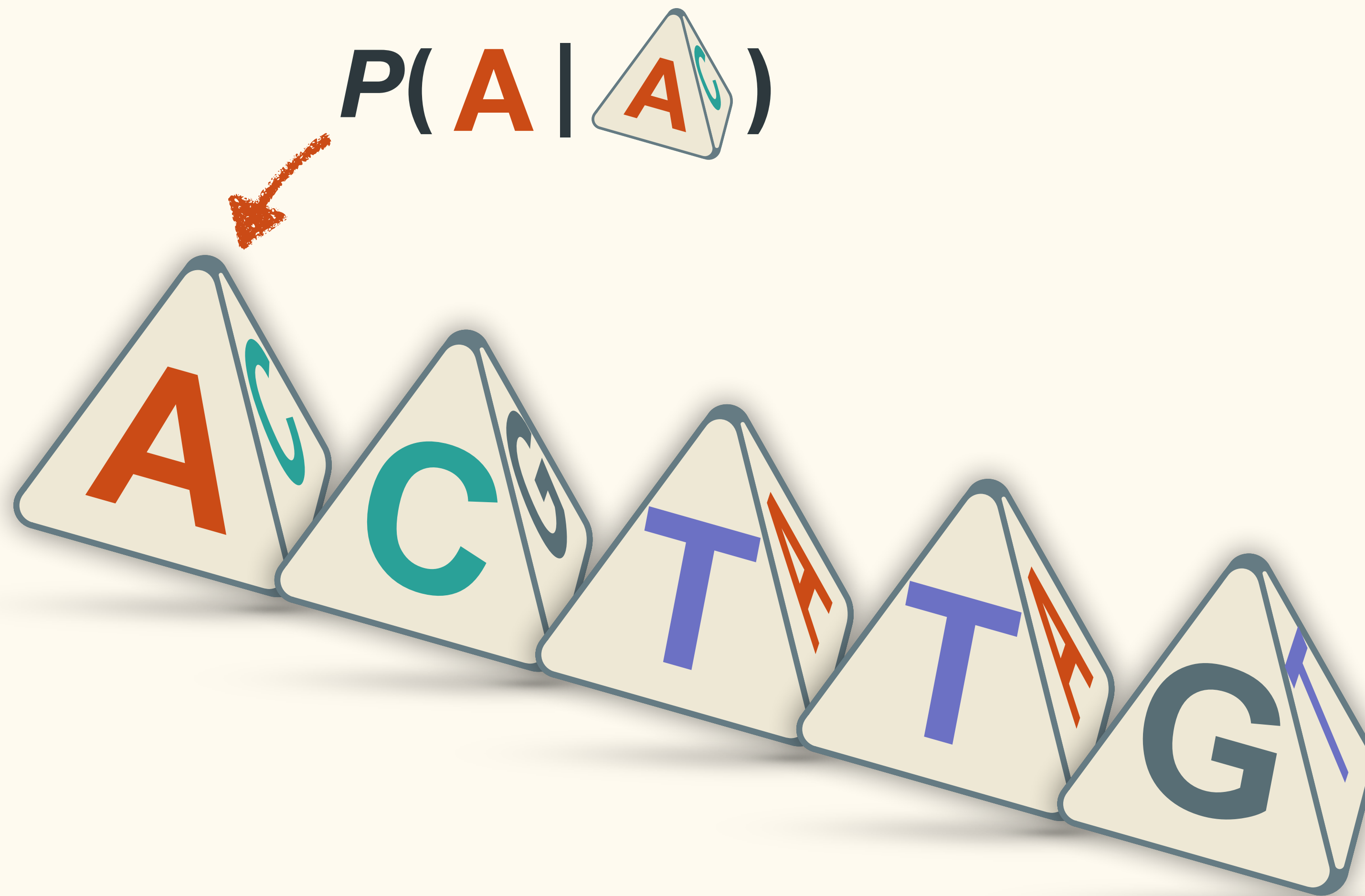
Probability

A, C, G, T

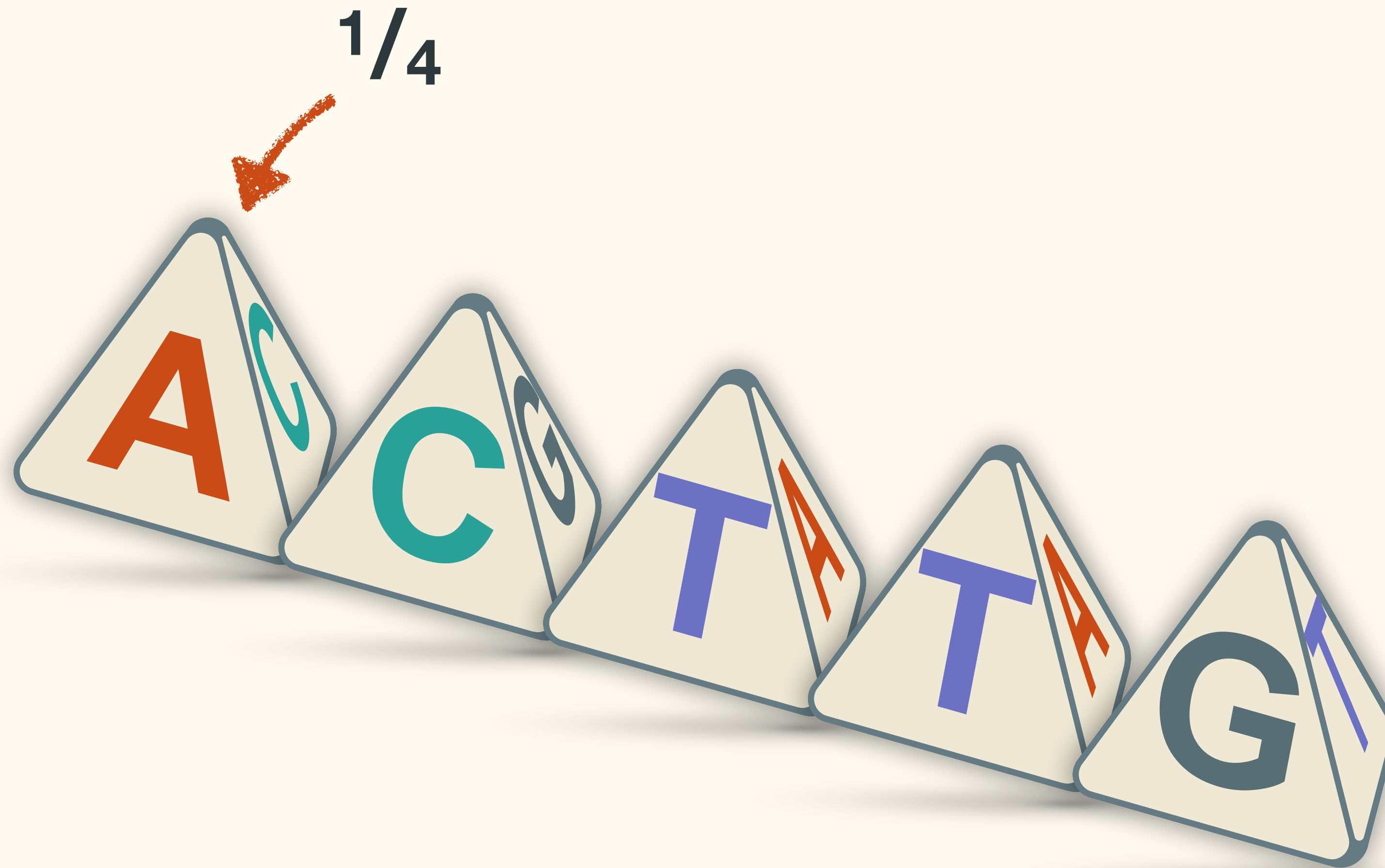
$$P(\text{A \& C \& T \& T \& G} \mid \text{AC}) = ?$$



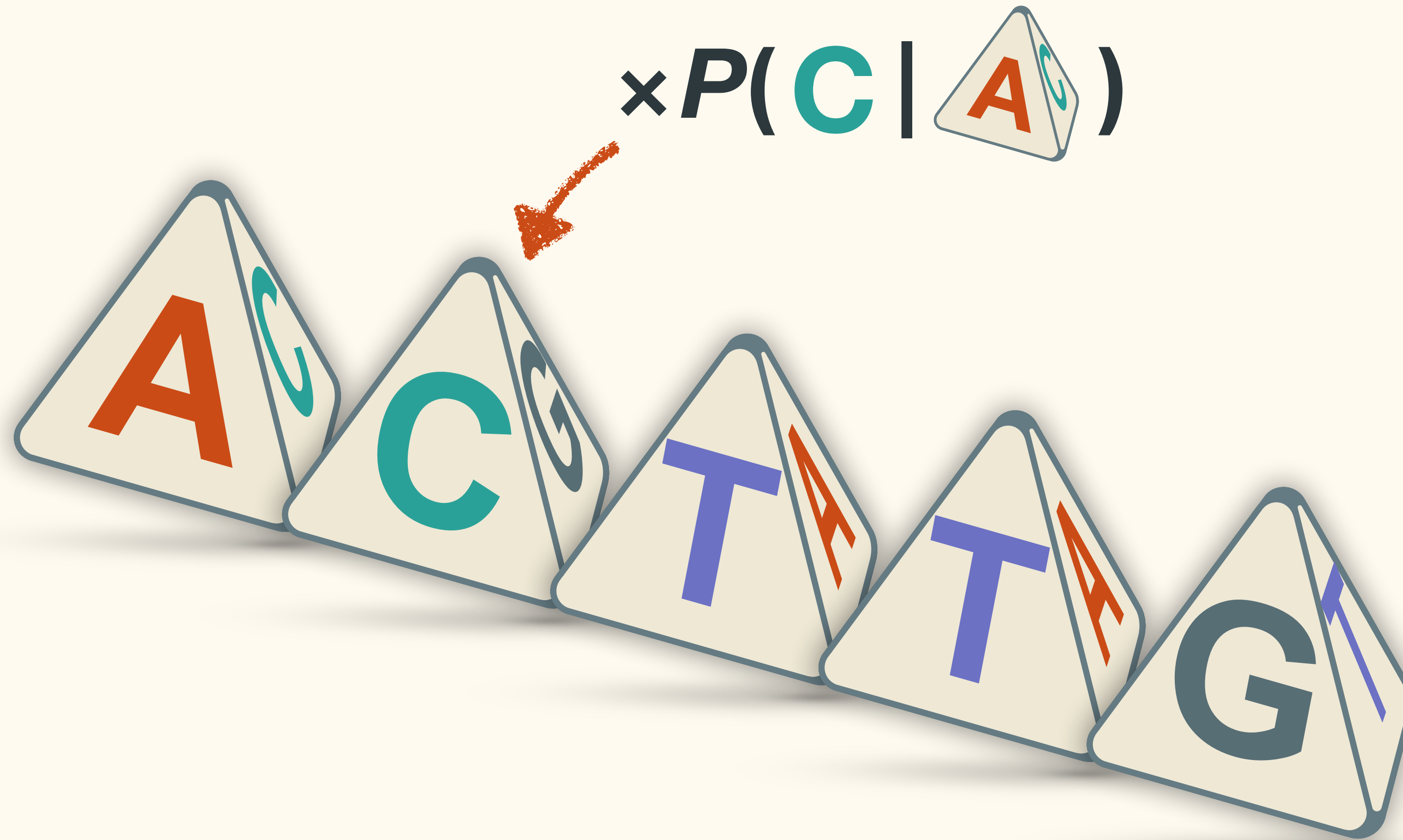
Probability



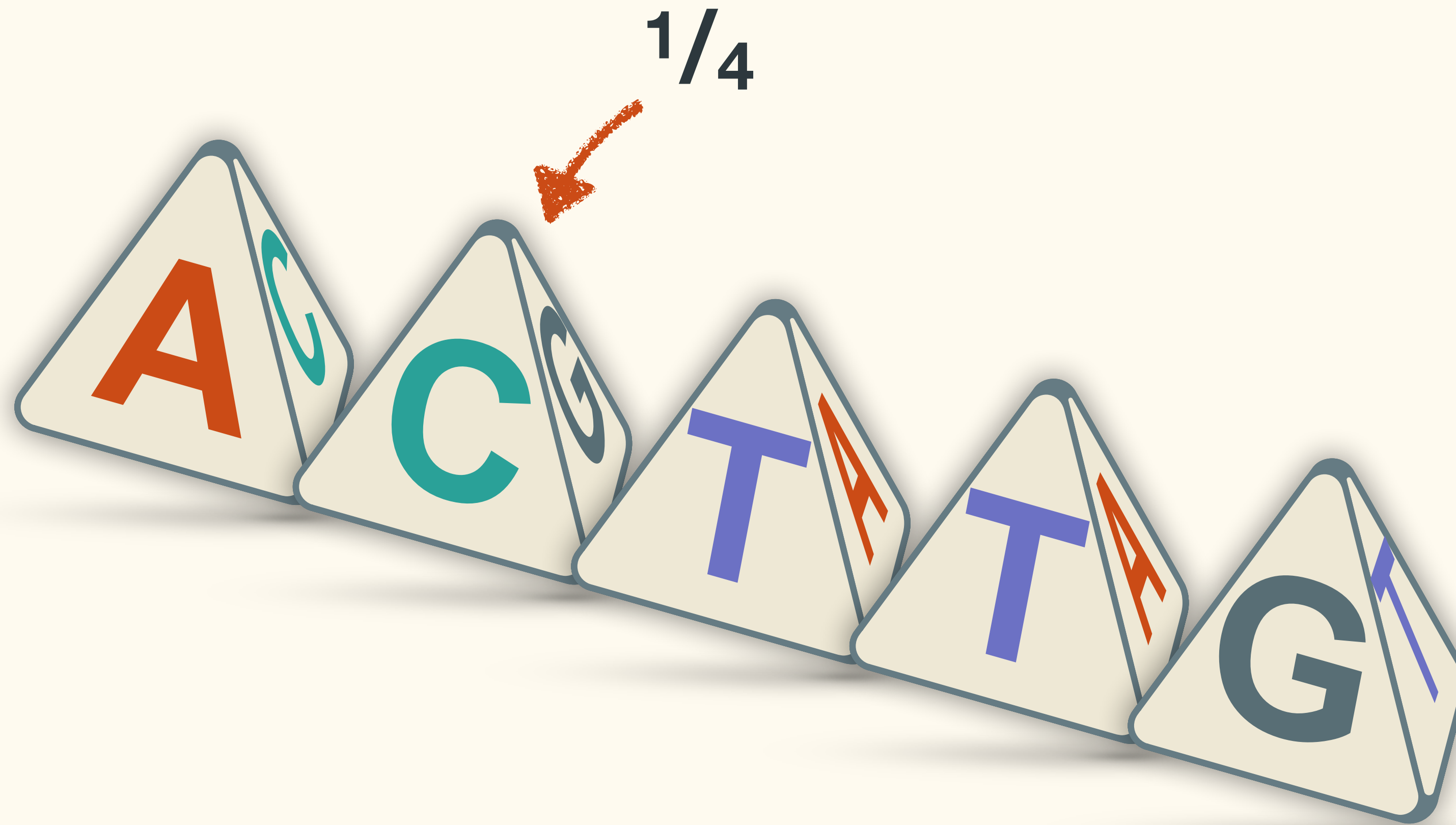
Probability



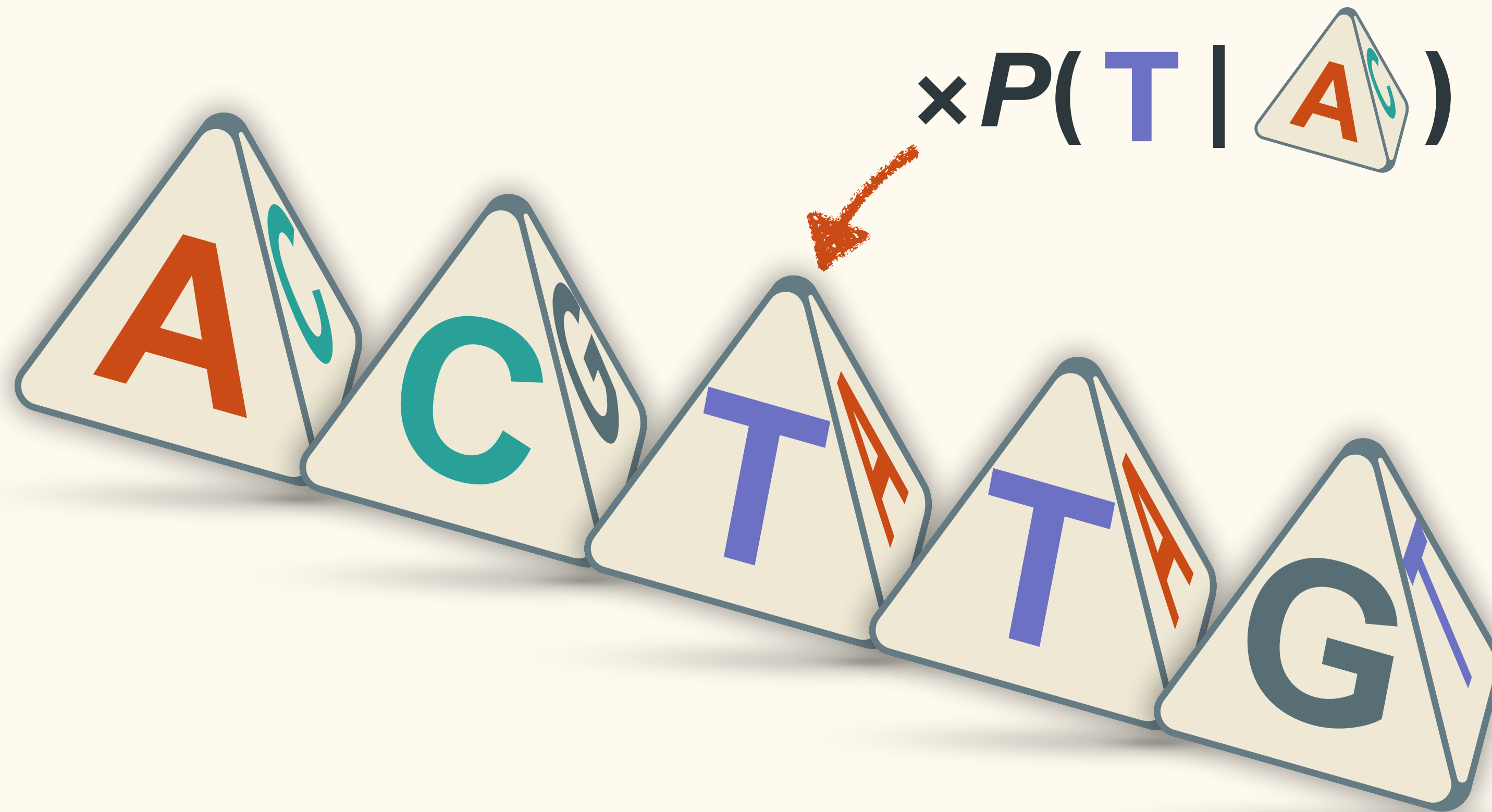
Probability



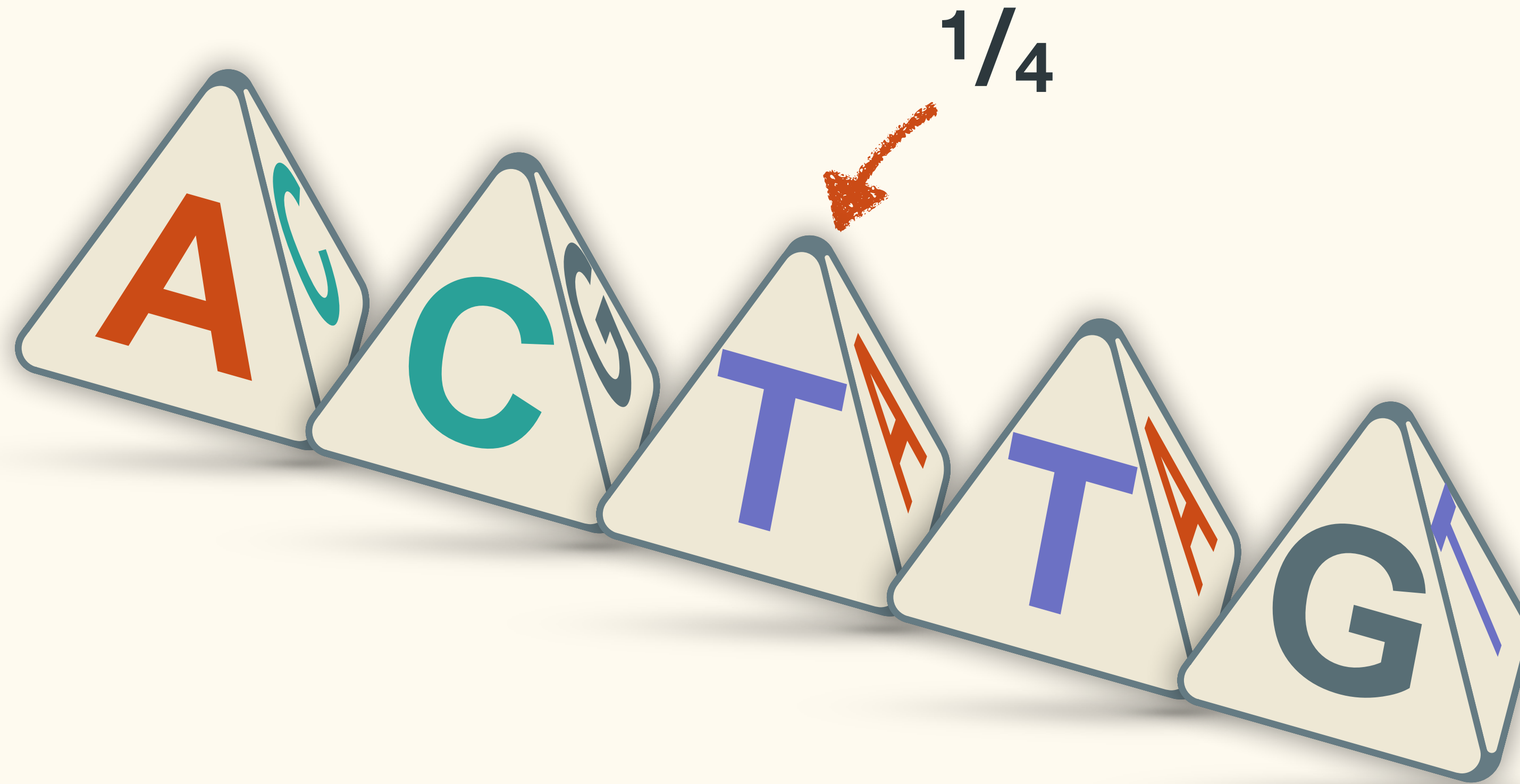
Probability



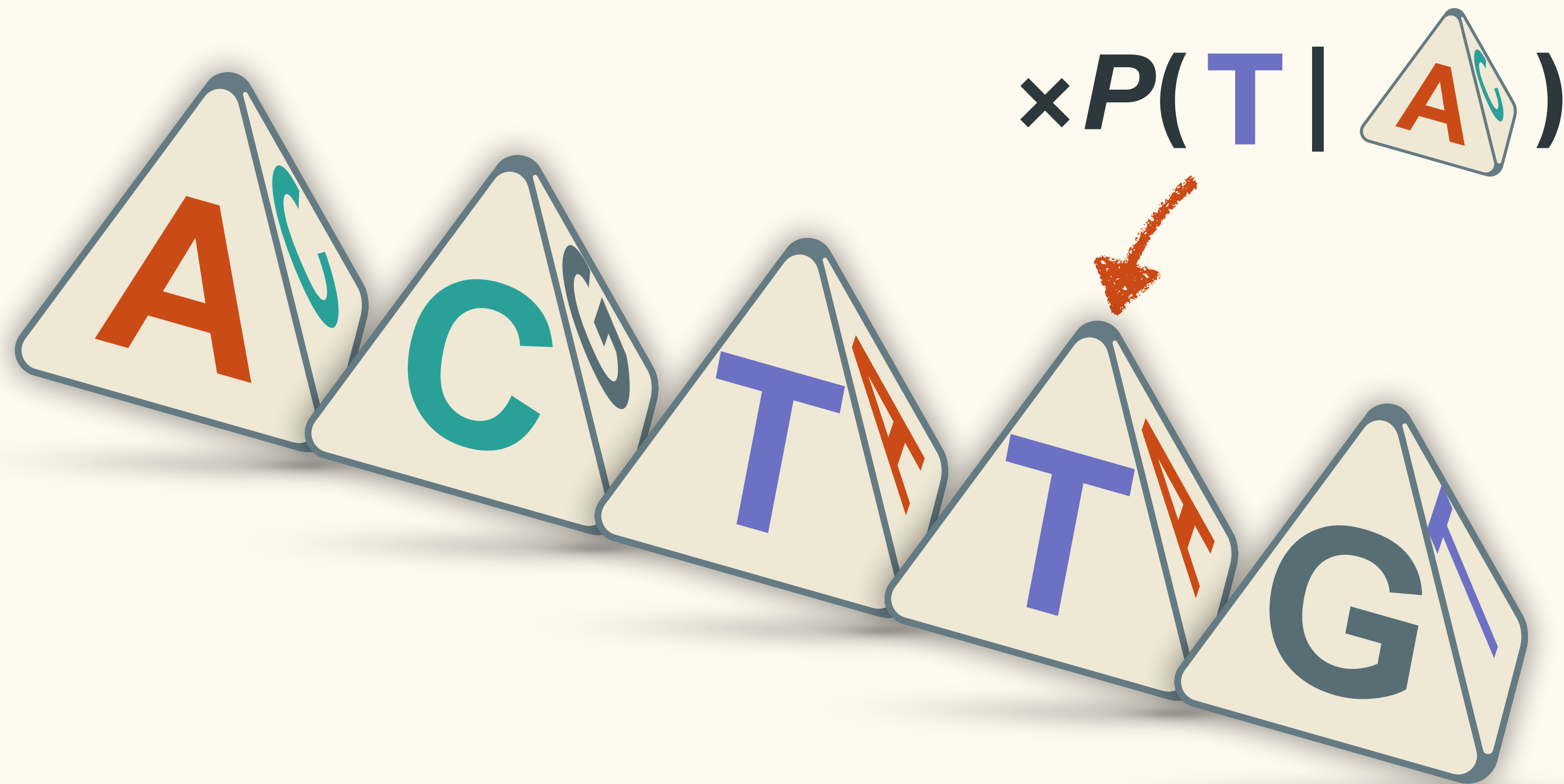
Probability



Probability



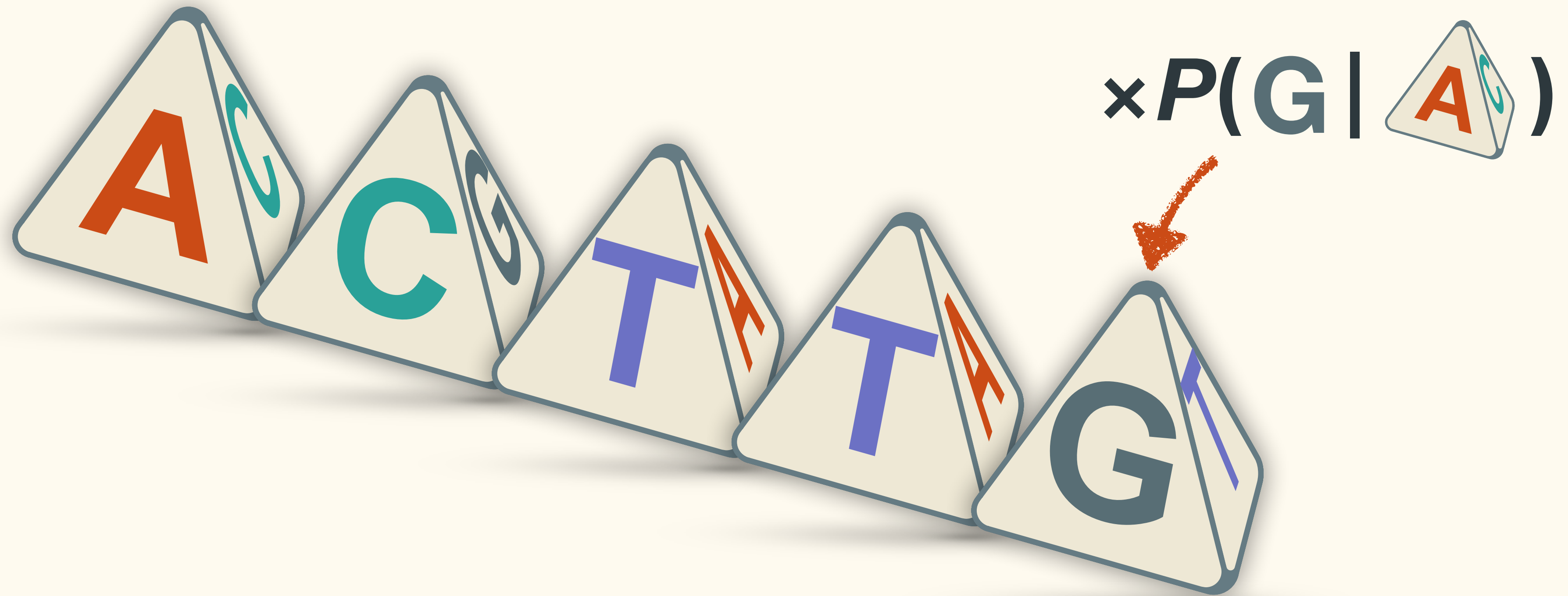
Probability



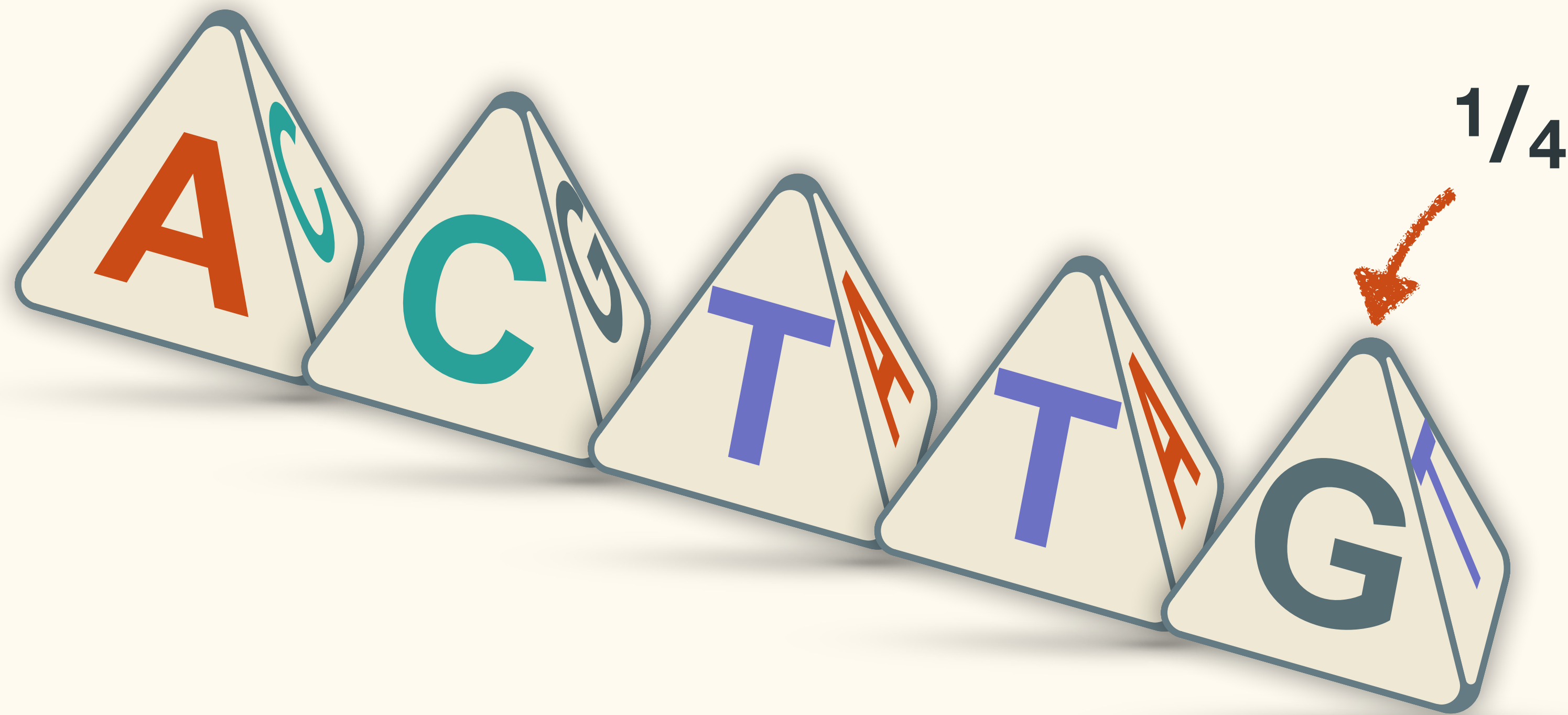
Probability



Probability

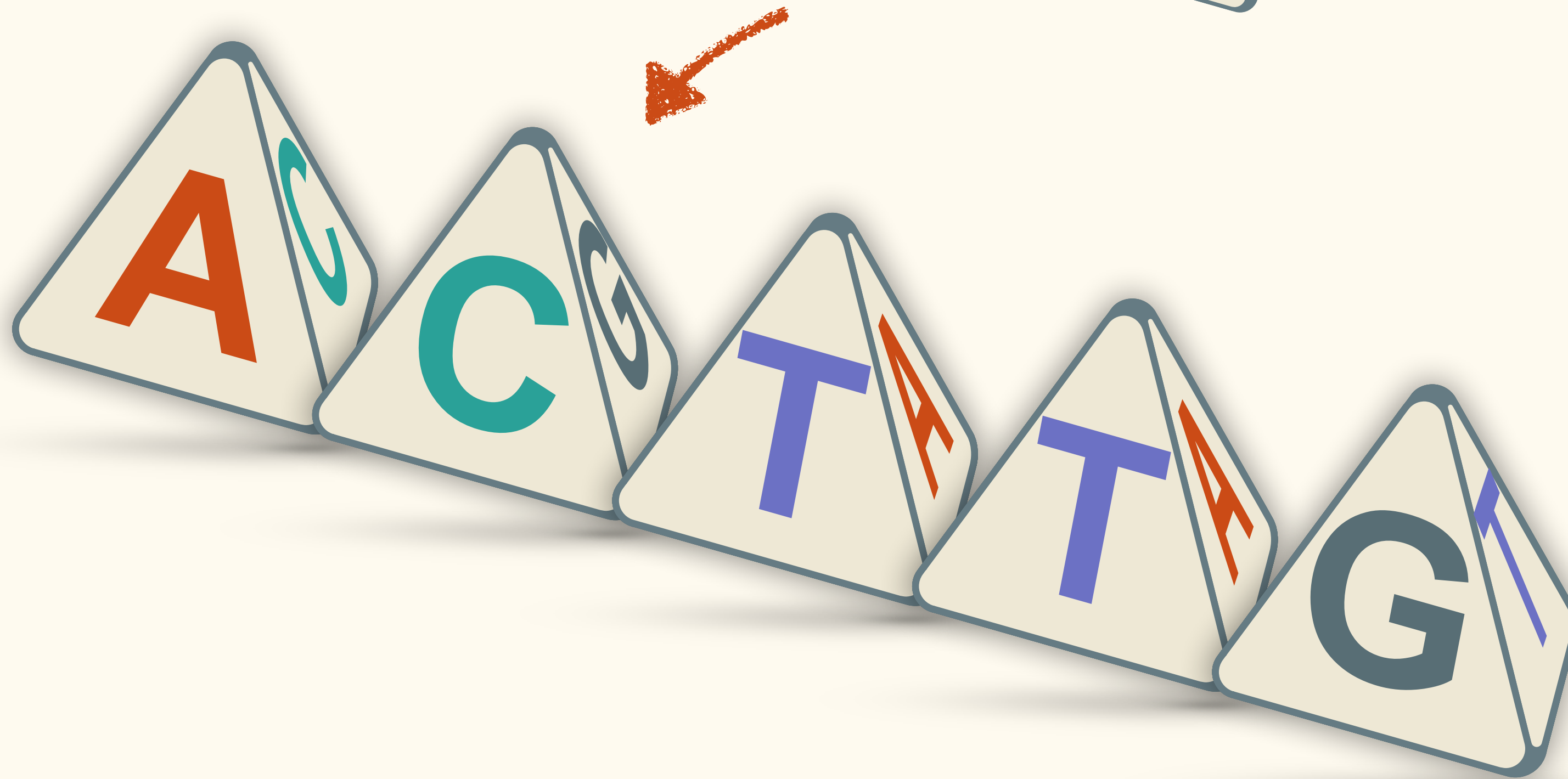


Probability

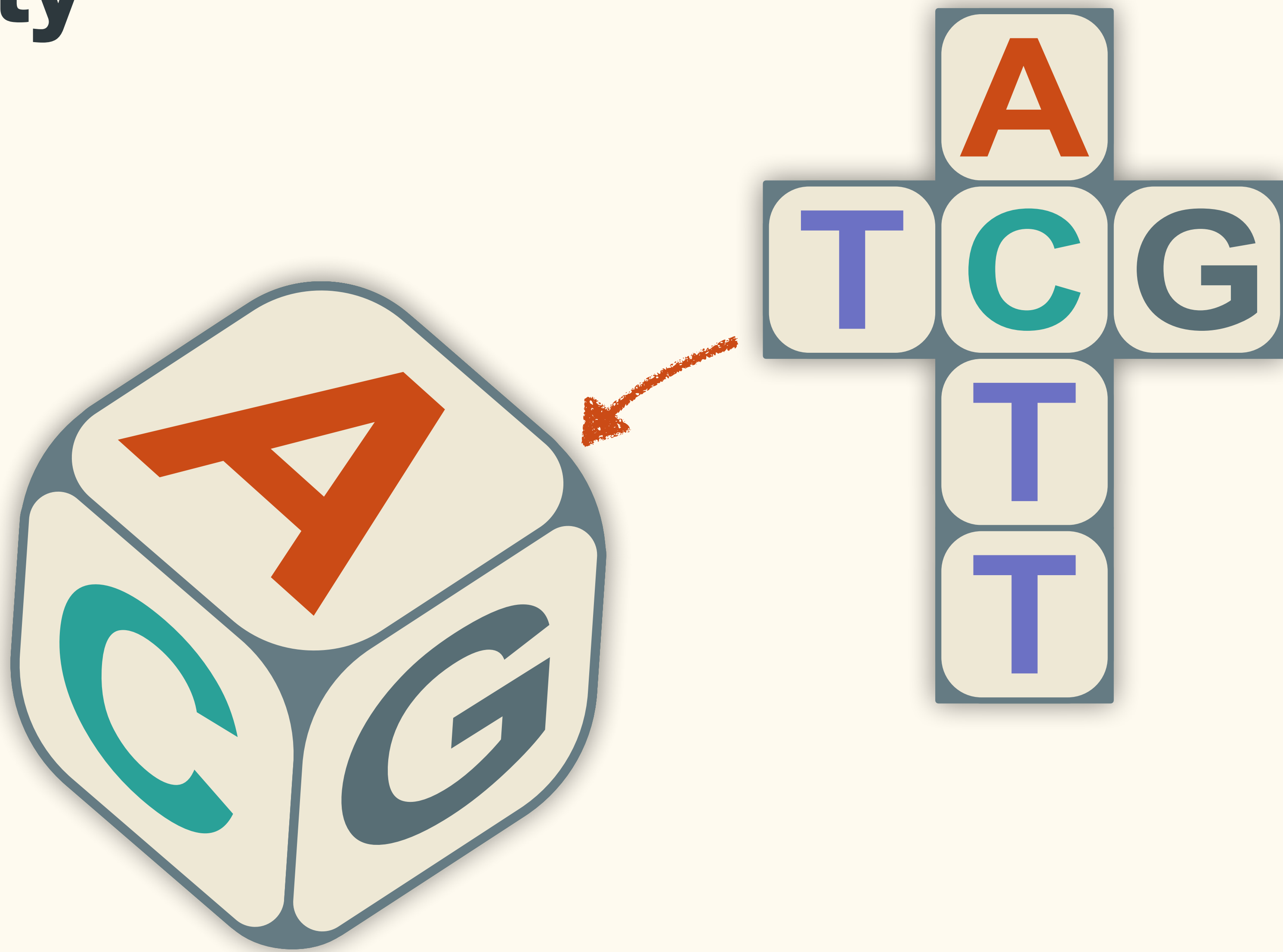


Probability

$$P(\text{A \& C \& T \& T \& G} \mid \text{AC}) = (1/4)^5$$

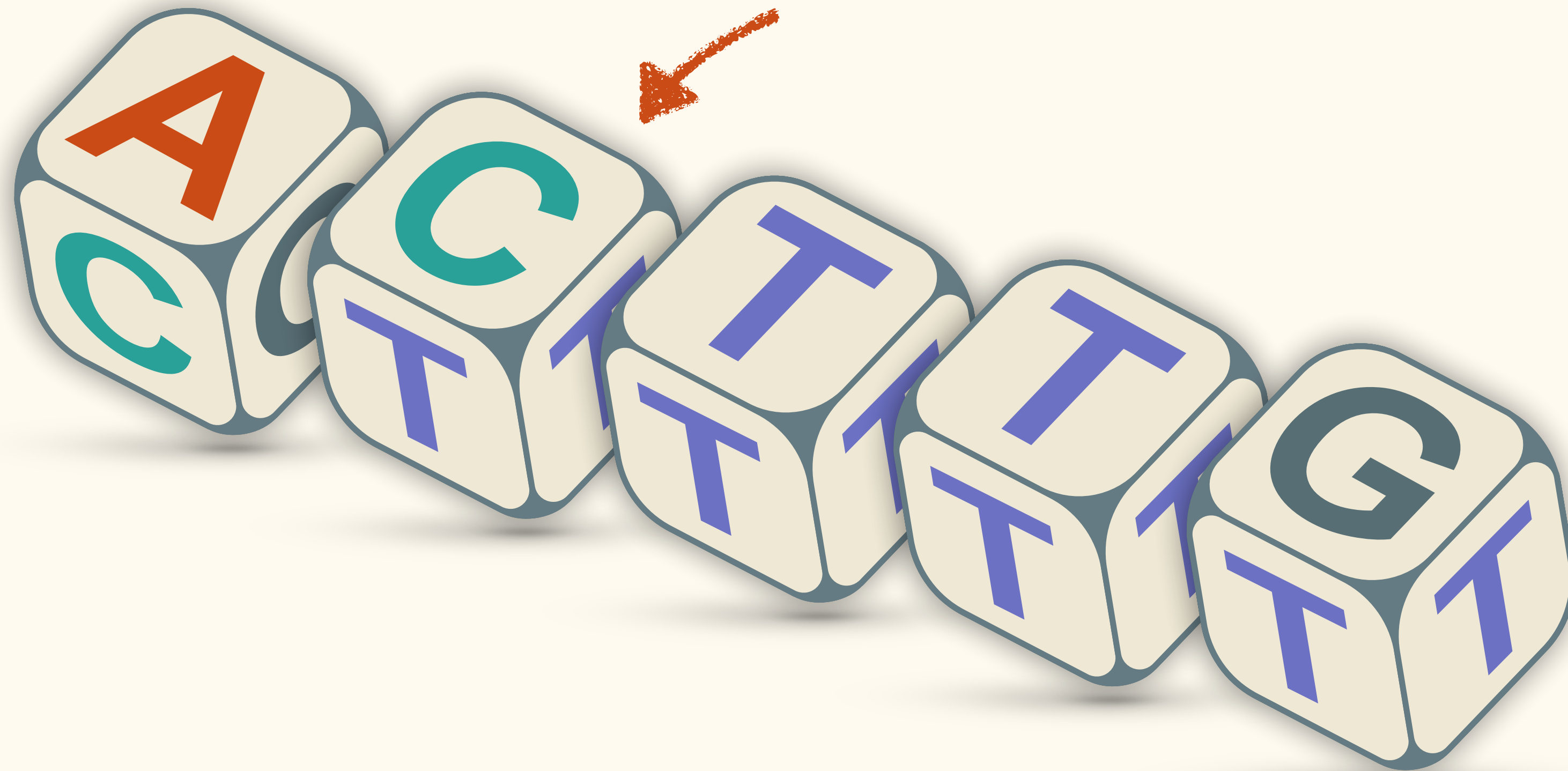


Probability



Probability

$$P(\text{ACTTG}) = ?$$

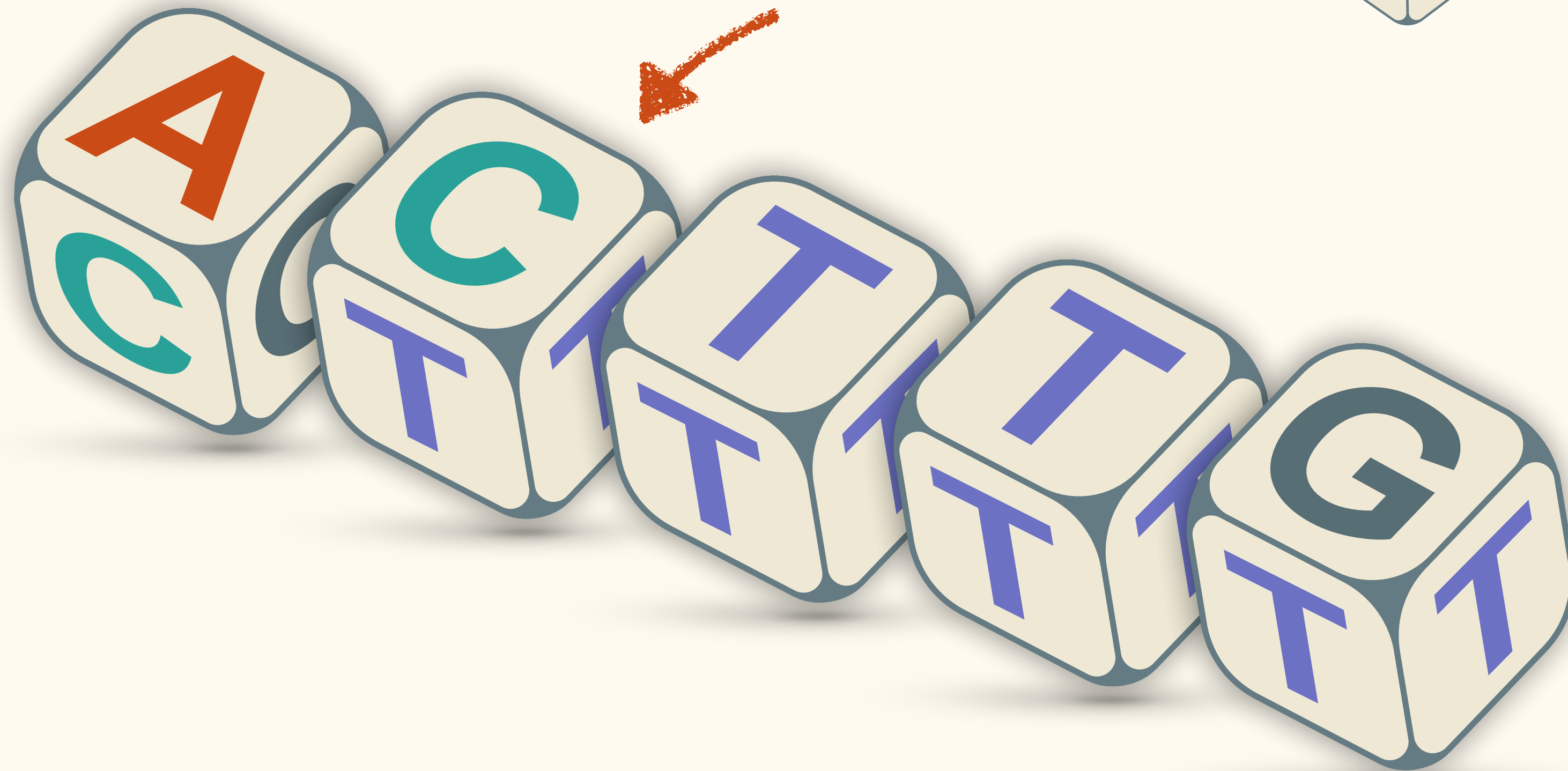


Probability

$f(T) > f(A), f(C), f(G)$



$$P(\text{ACTTGG} \mid \text{die}) = ?$$

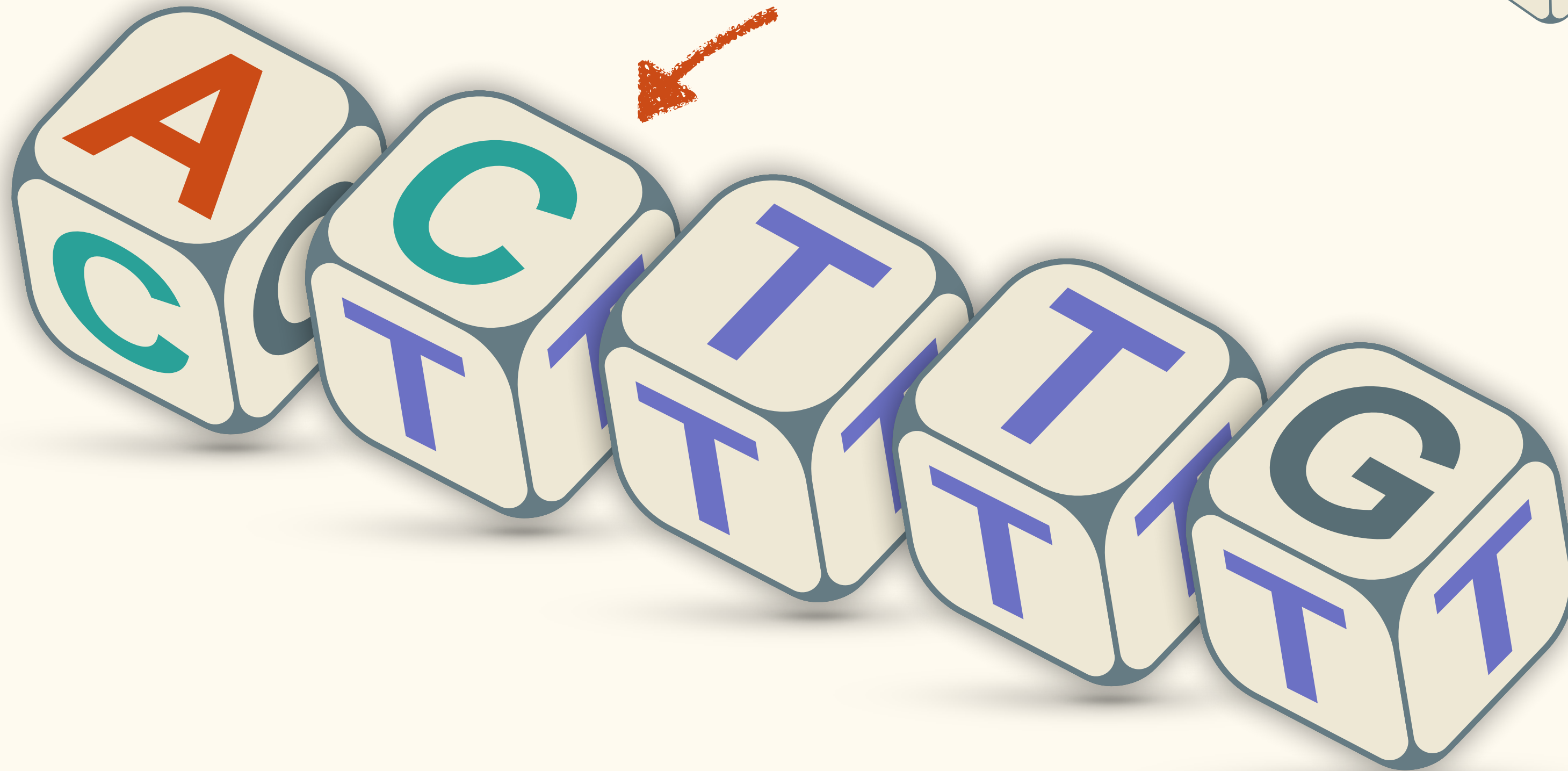


Probability

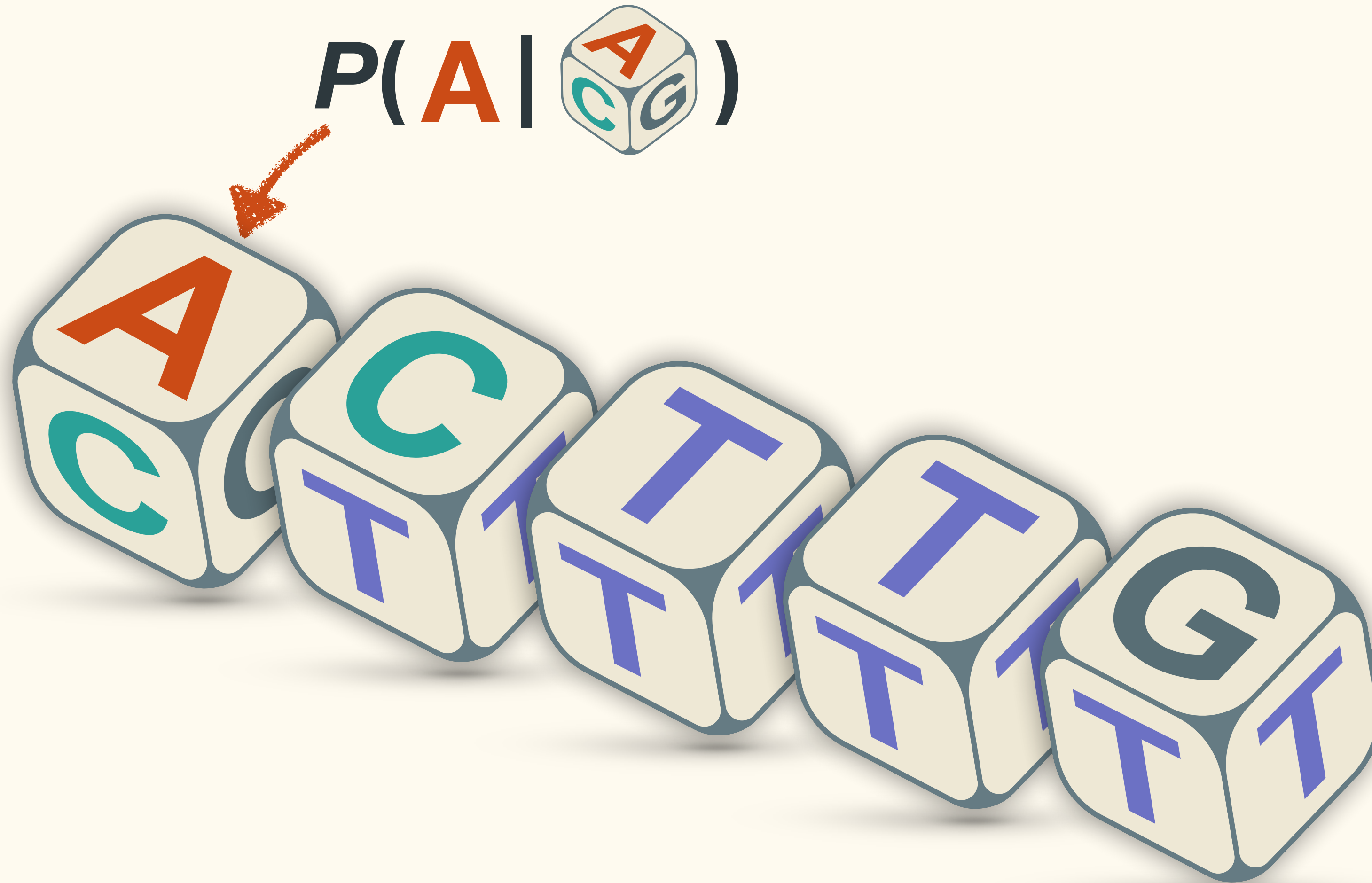
$f(T) > f(A), f(C), f(G)$



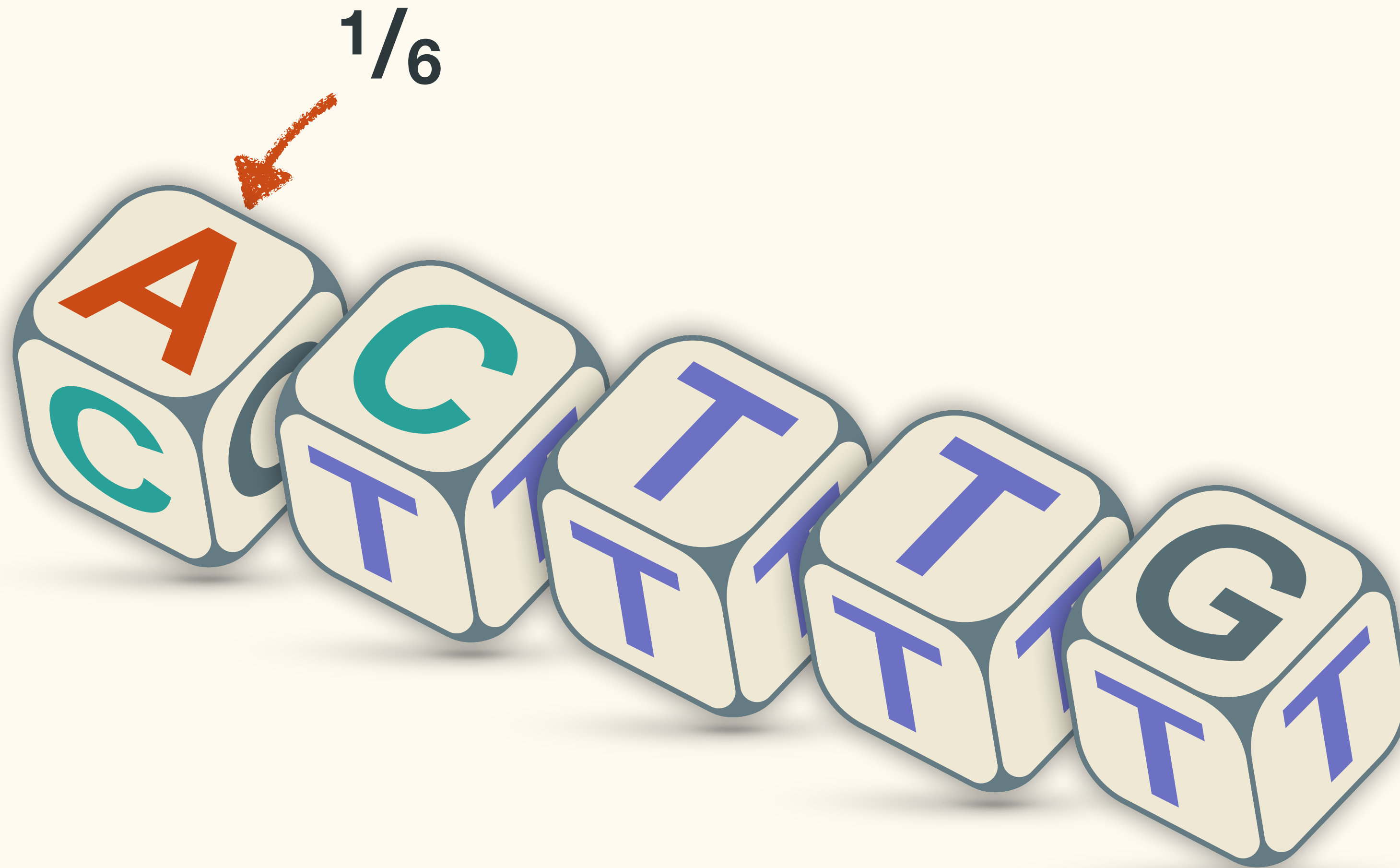
$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = ?$$



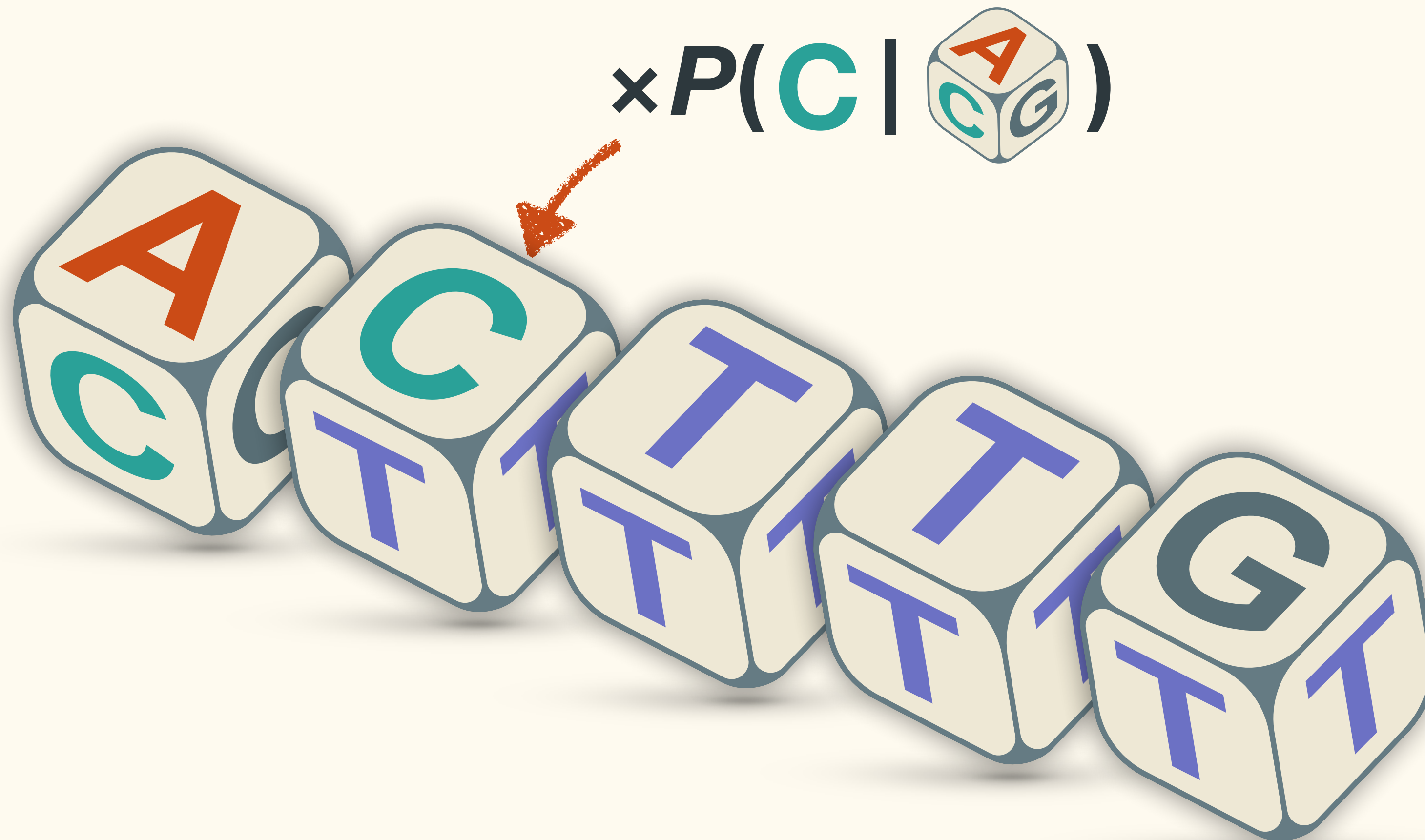
Probability



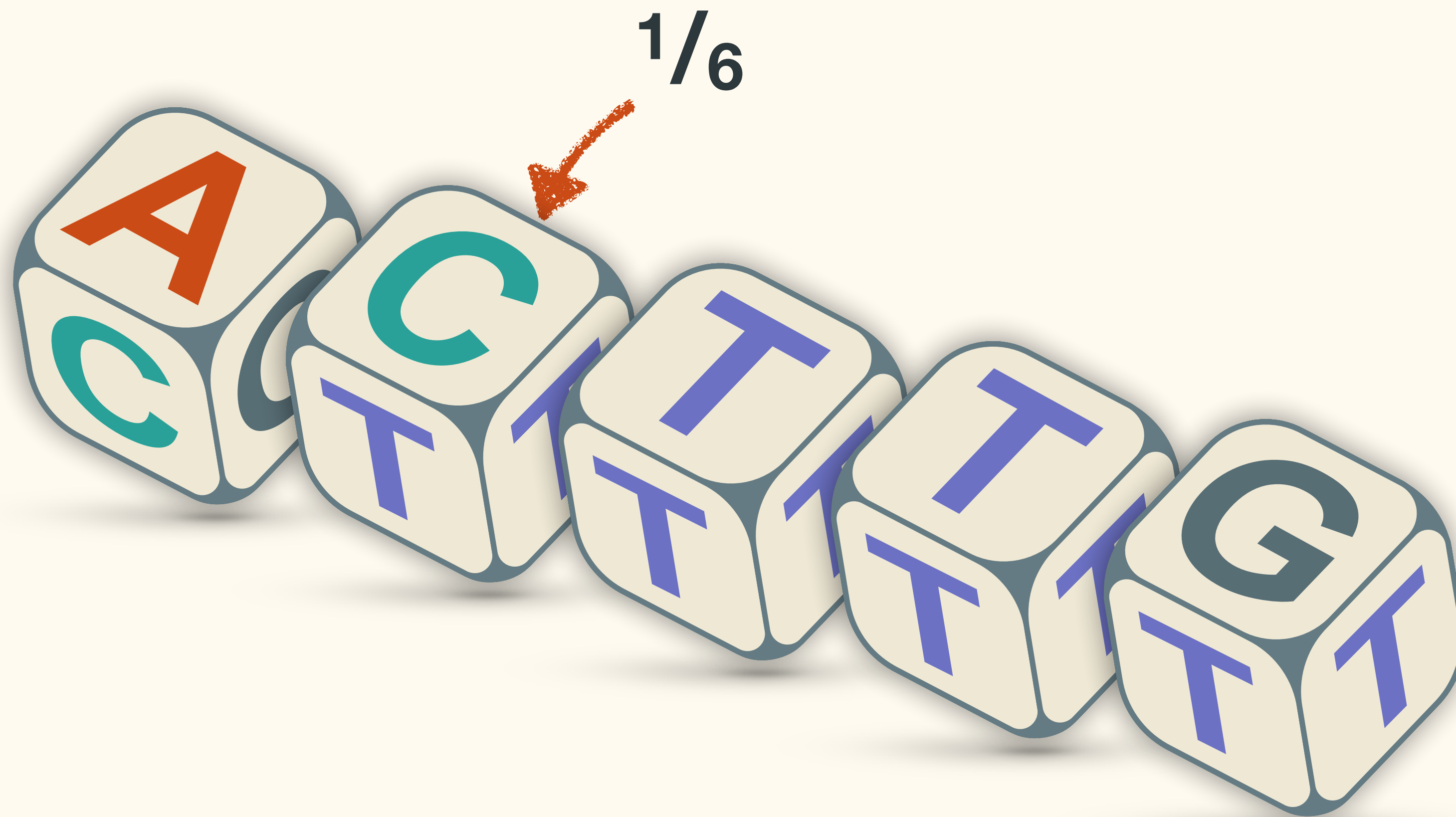
Probability



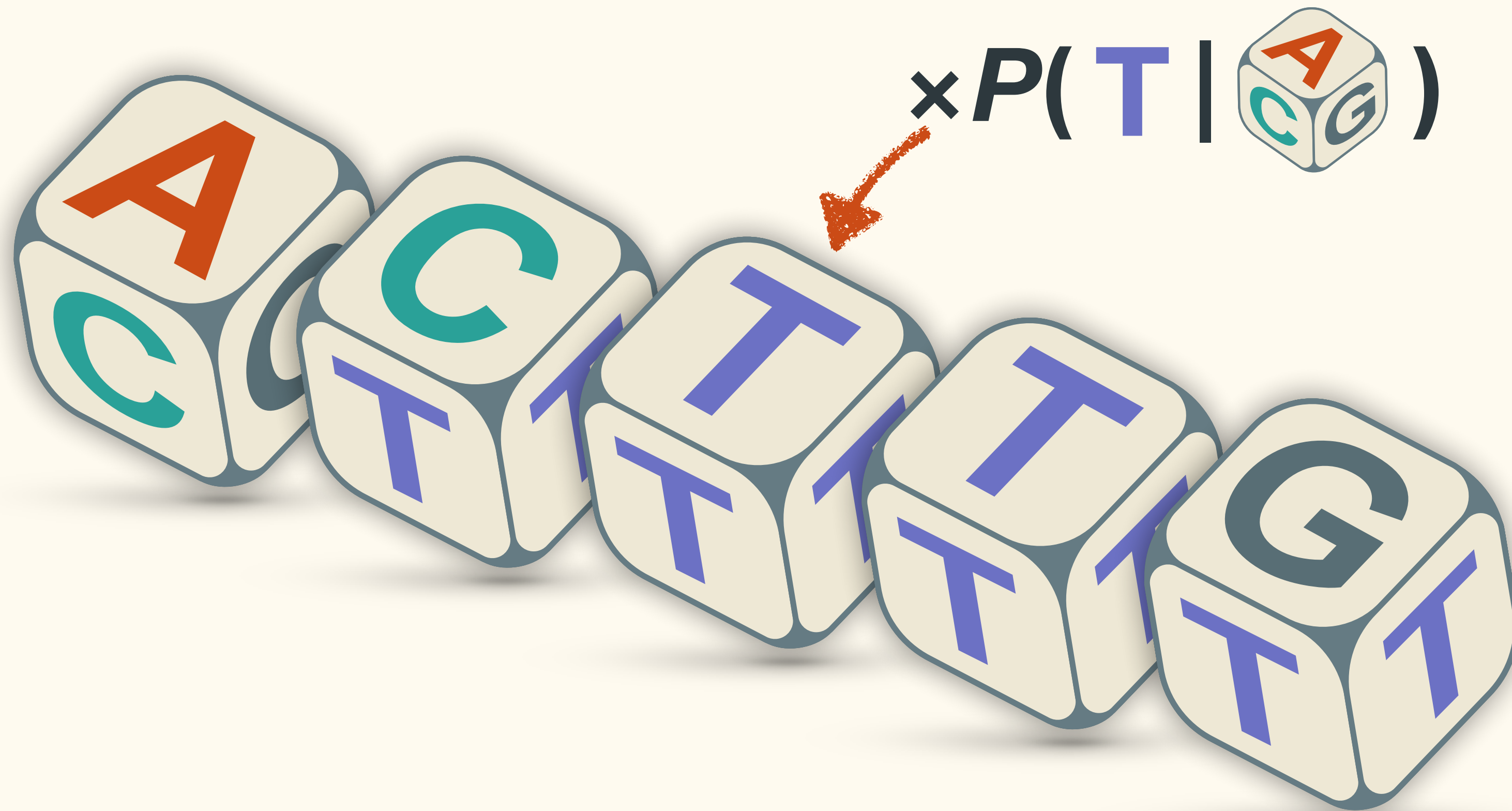
Probability



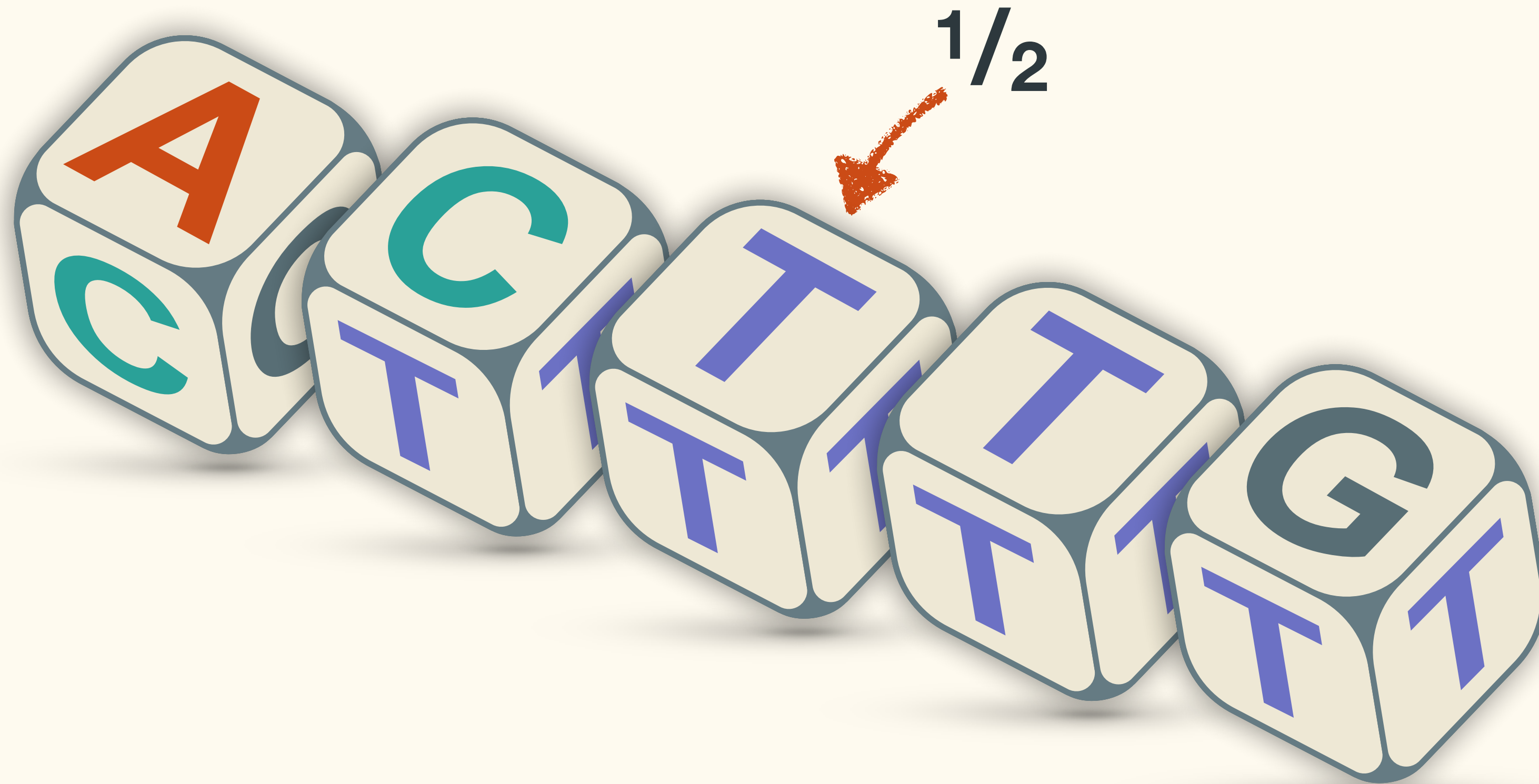
Probability



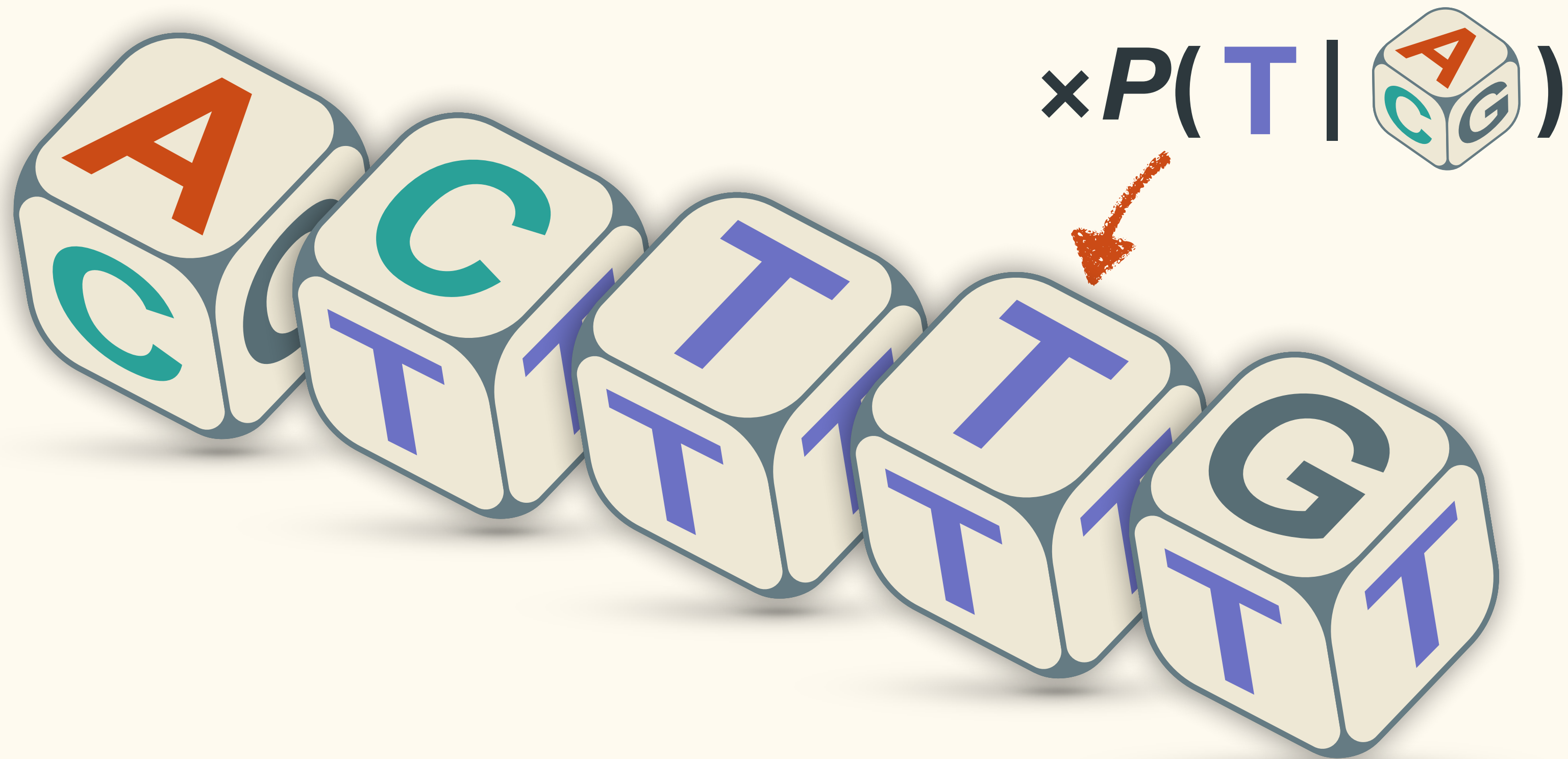
Probability



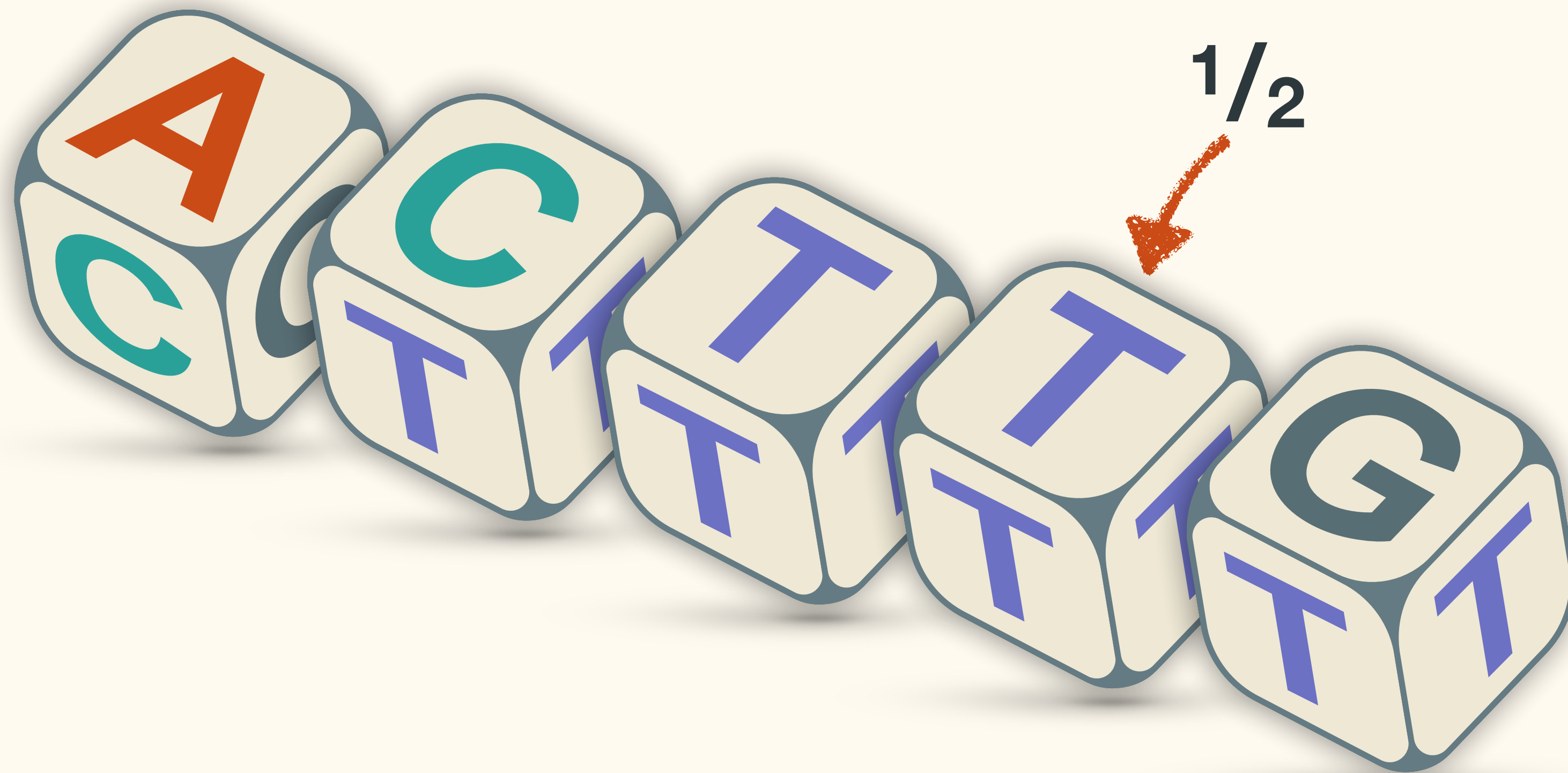
Probability



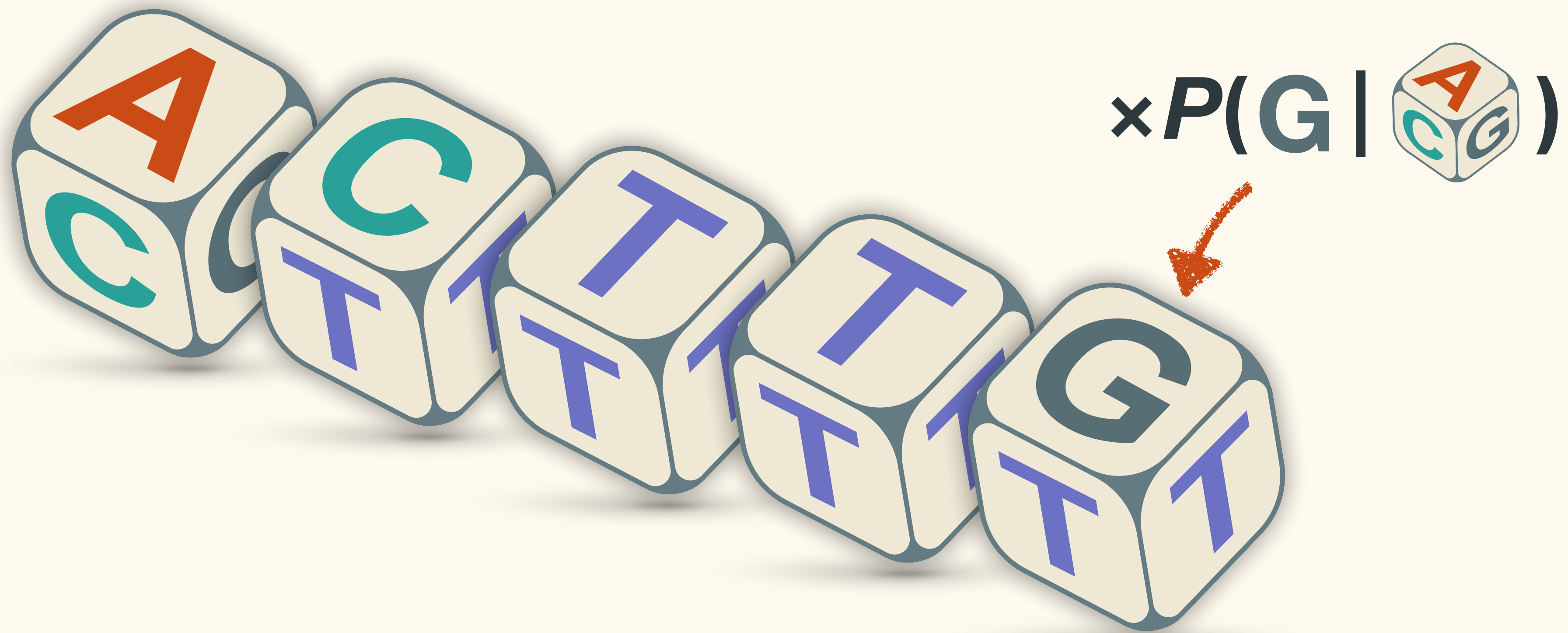
Probability



Probability



Probability

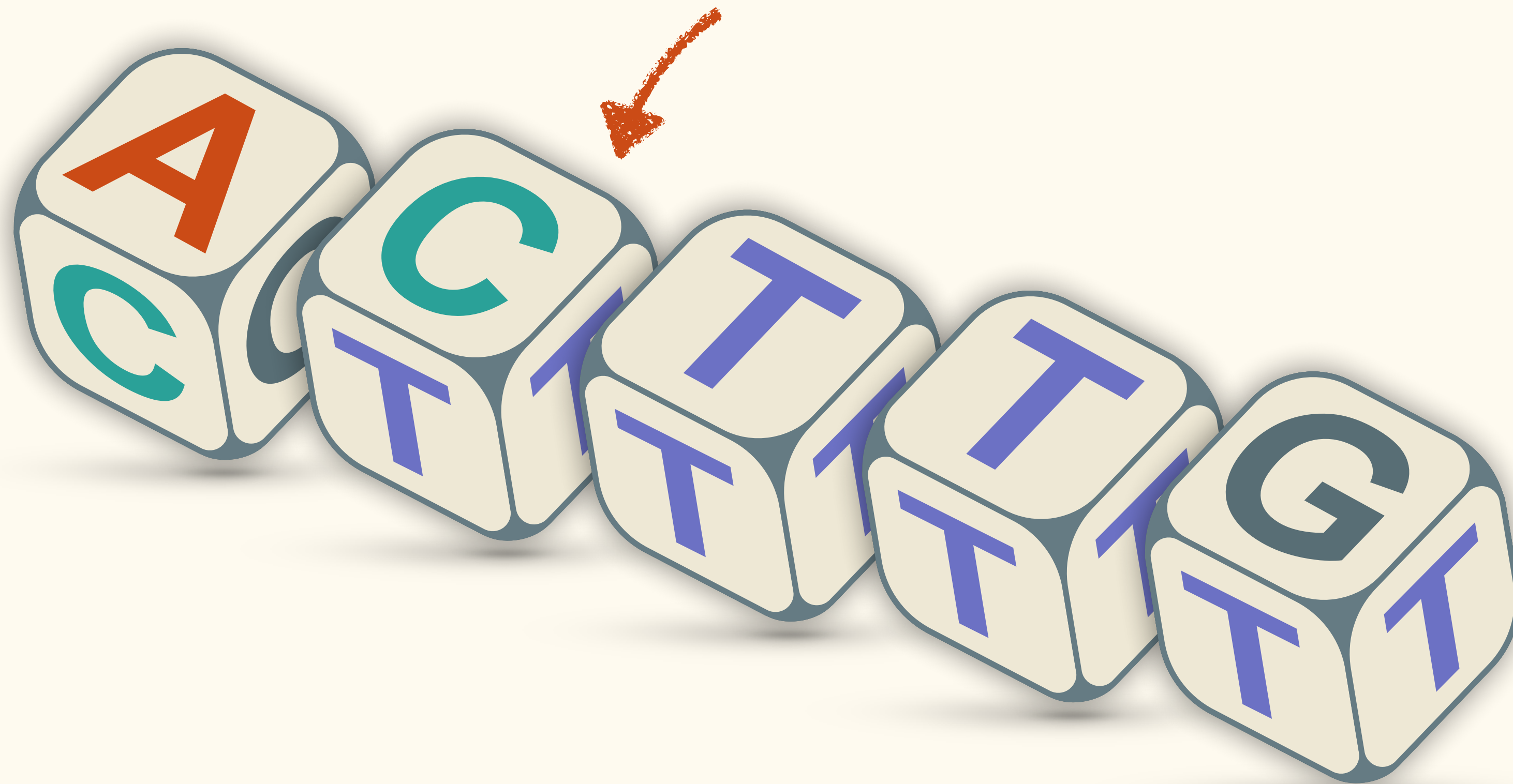


Probability



Probability

$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = ?$$



Probability

$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = \left(\frac{1}{6}\right)^3 \left(\frac{1}{2}\right)^2$$



Probability

$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = \left(\frac{1}{4}\right)^5$$

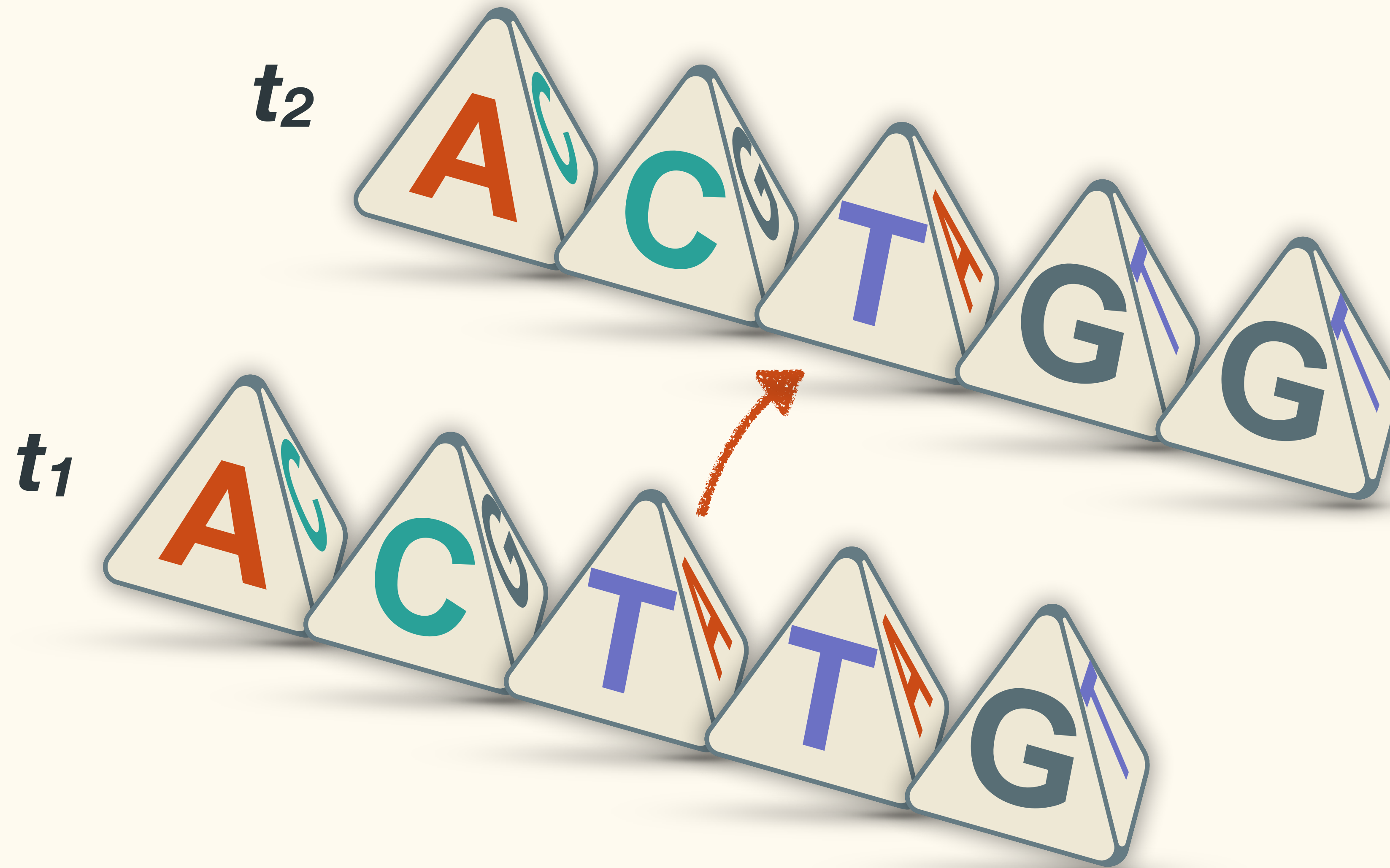
$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = \left(\frac{1}{6}\right)^3 \left(\frac{1}{2}\right)^2$$

Probability

$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = 0.0010$$

$$P(\mathbf{A} \& \mathbf{C} \& \mathbf{T} \& \mathbf{T} \& \mathbf{G} \mid \text{die}) = 0.0012$$

Probability



Probability

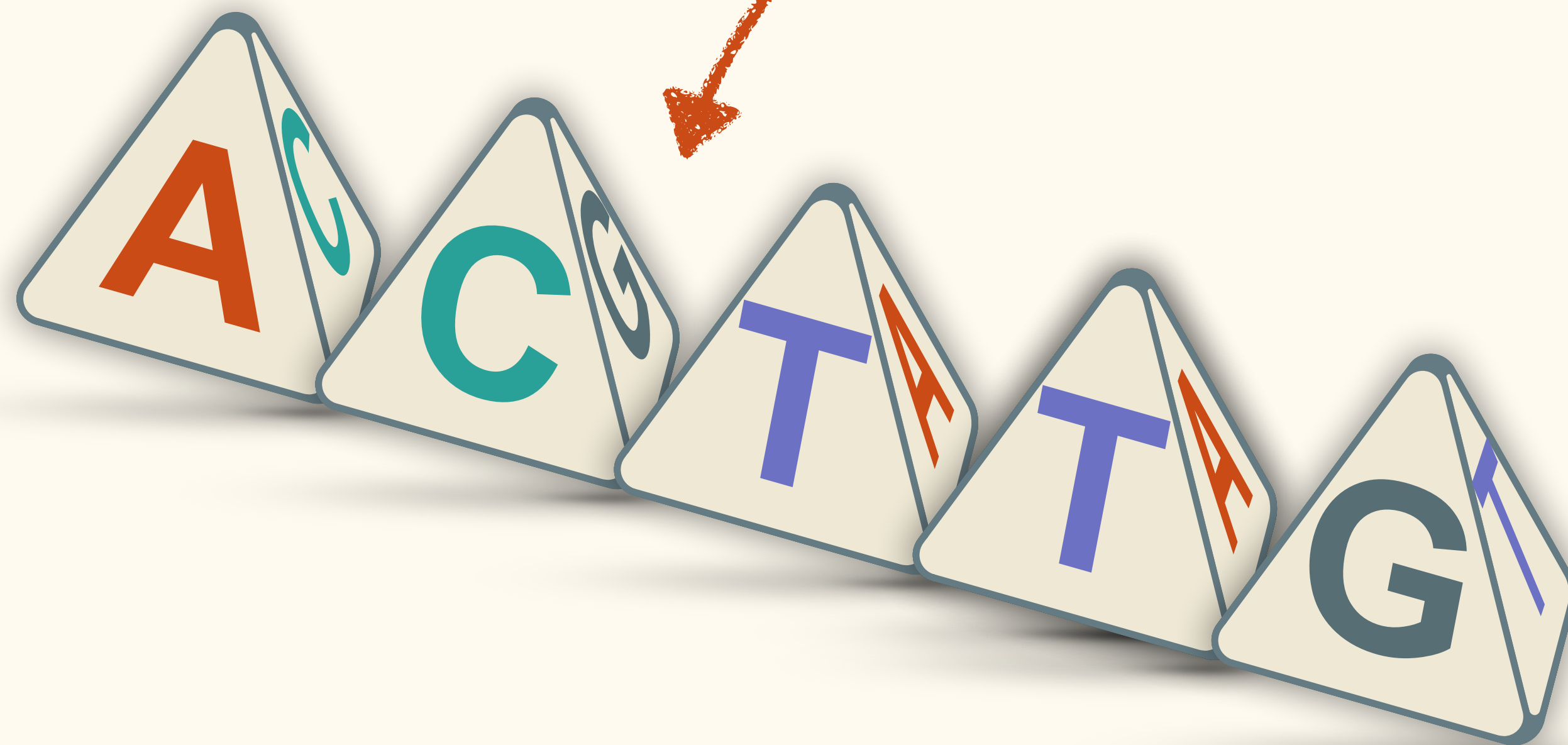
$$P\left(\begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array}\right) = ?$$



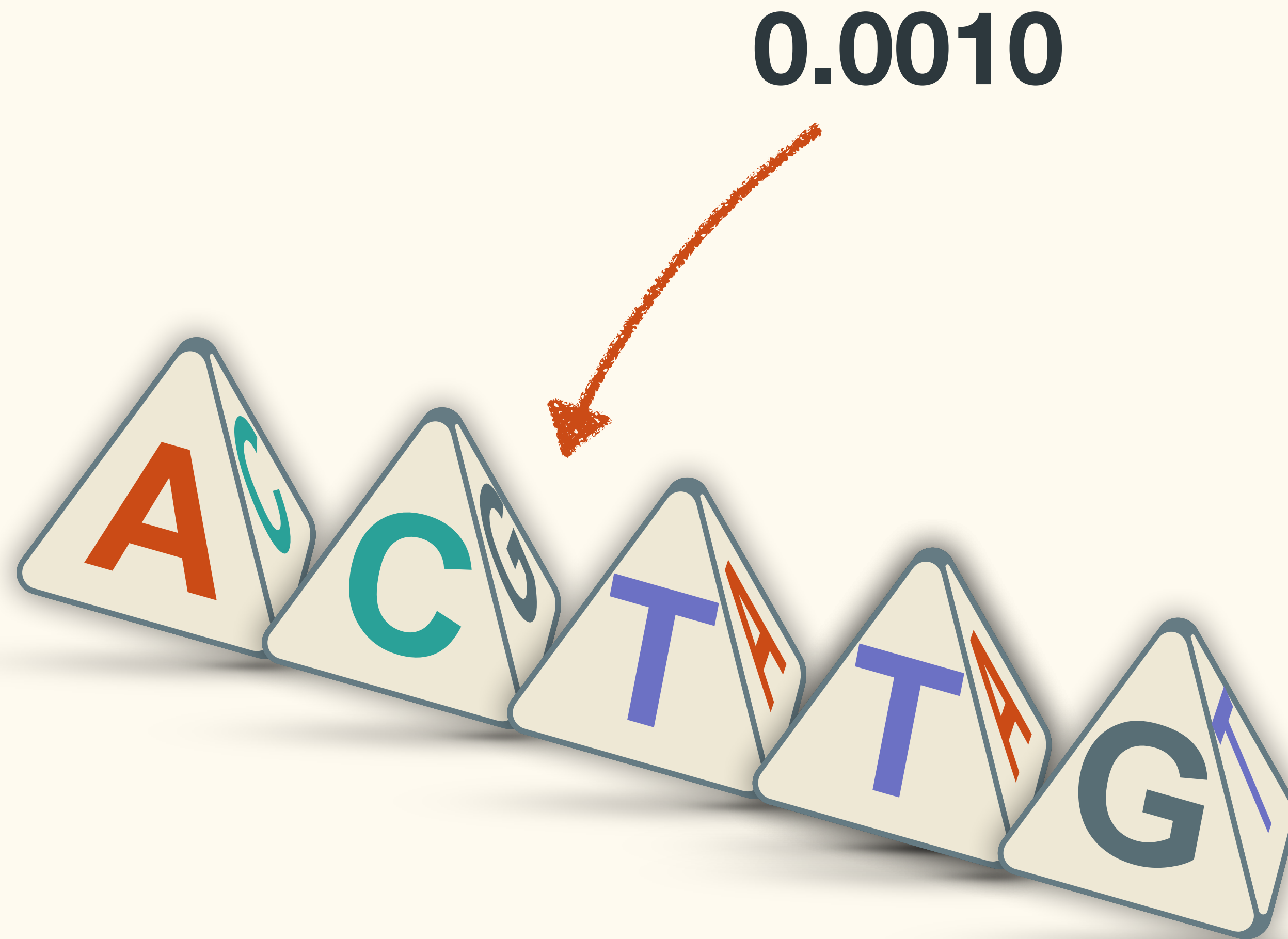
Probability

A, C, G, T

$$P(\text{ACTTG} \mid \triangle_{AC})$$



Probability



Probability

$\times P(\text{T} \rightarrow \text{G})$



Probability

$$\times P(\text{T} \rightarrow \text{G} \mid \text{AC})$$

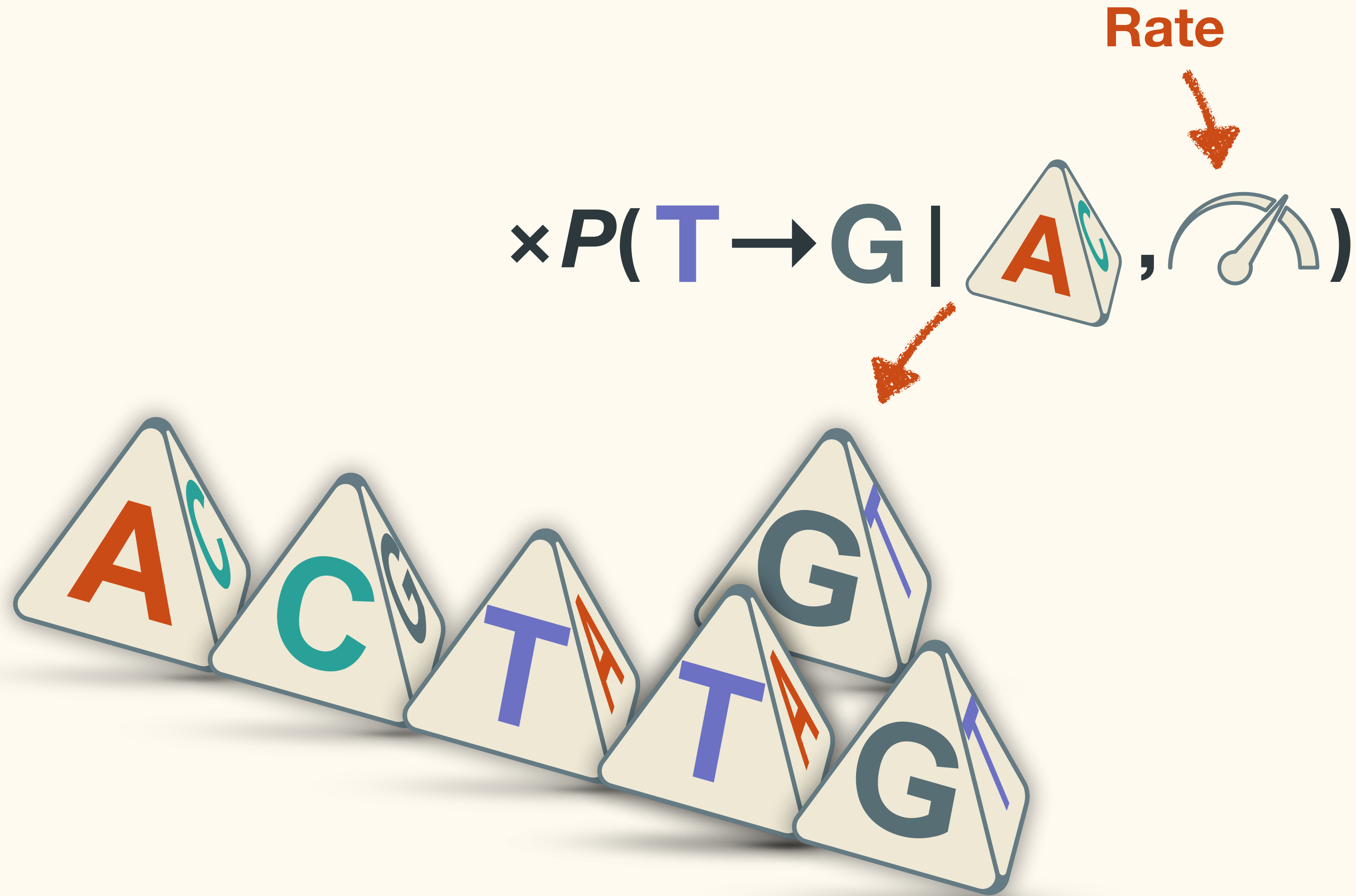


Probability

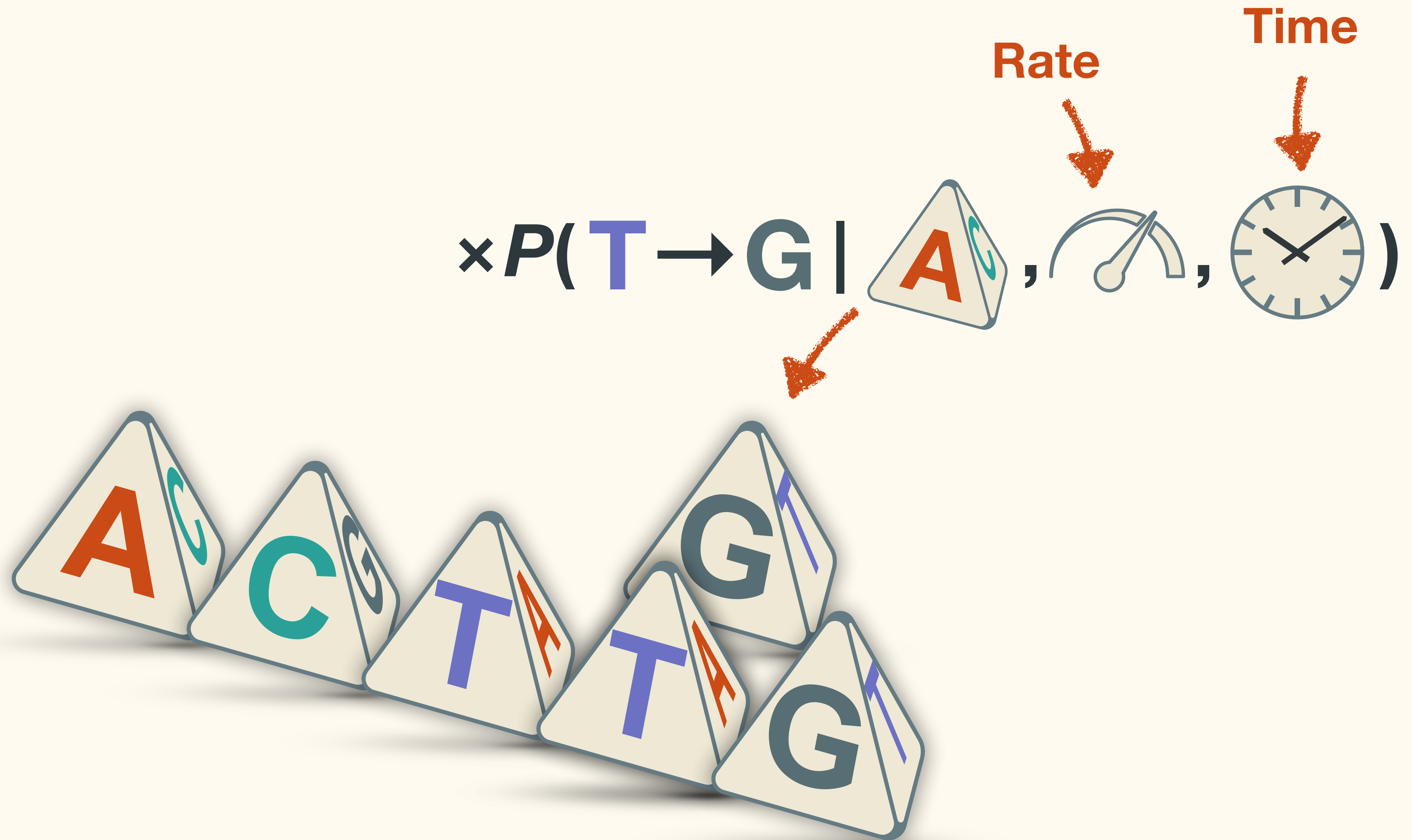
$\times P(\text{T} \rightarrow \text{G} \mid \triangle \text{AC}, ?)$



Probability



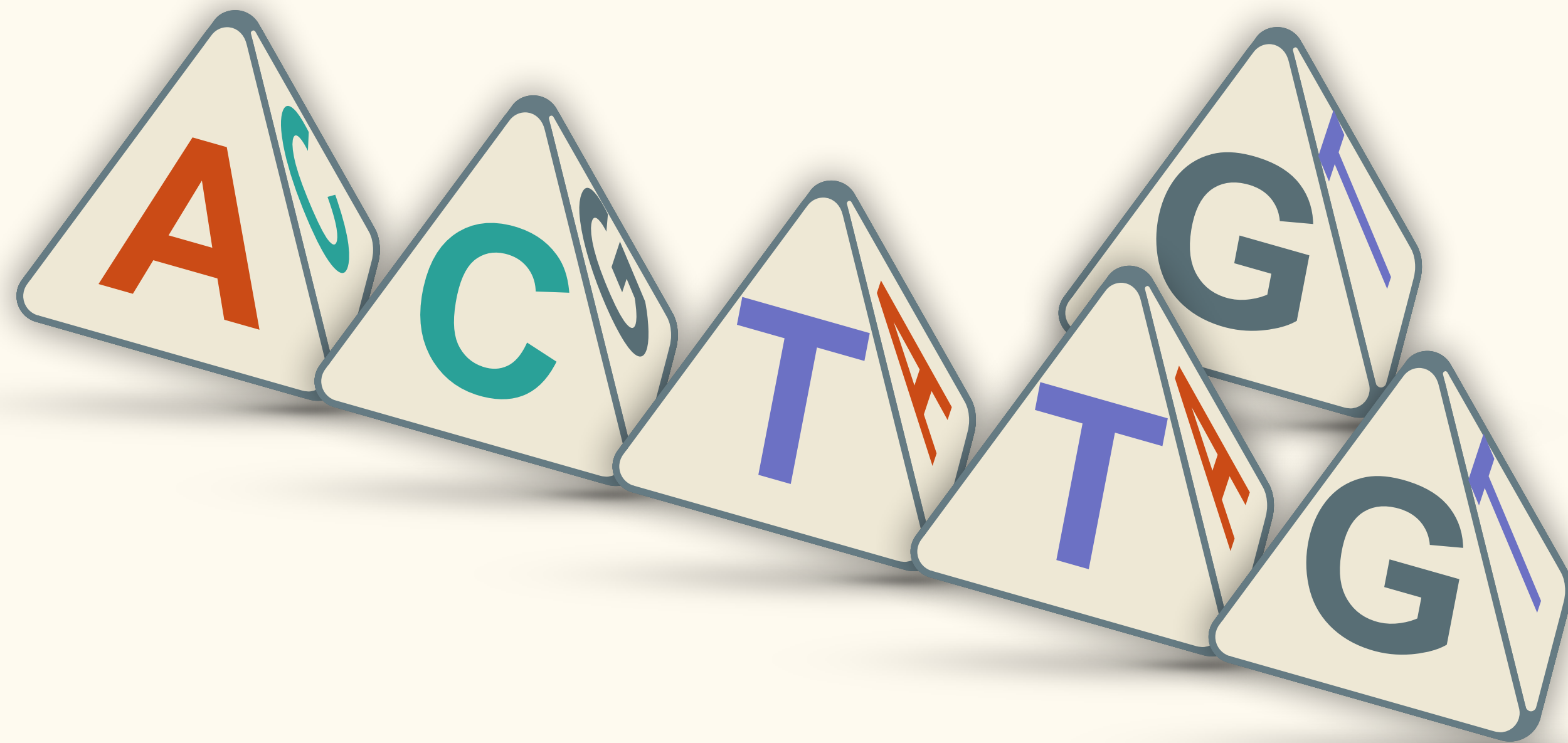
Probability



Probability

Rate Time

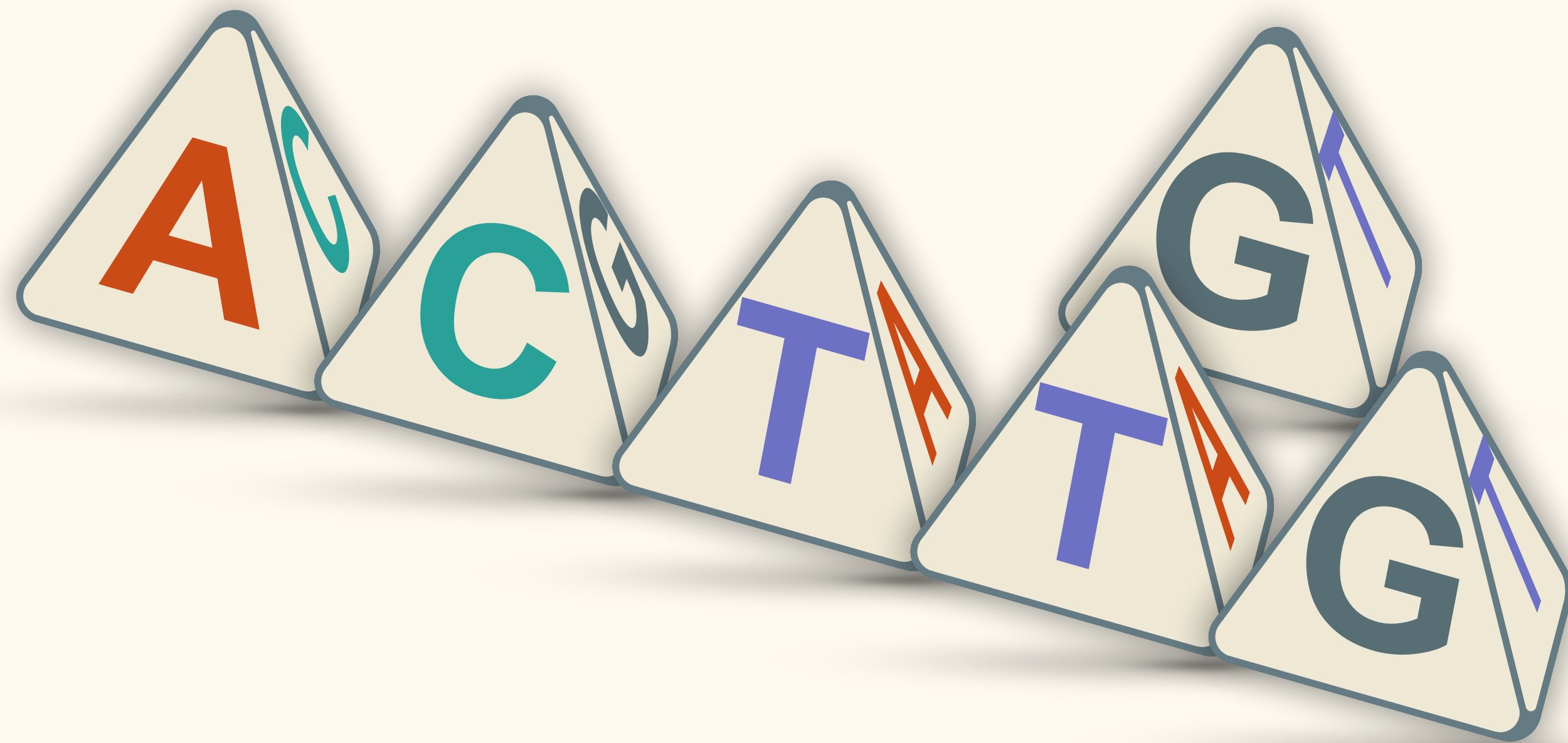
$$\left(\frac{1}{4}\right) \times \left(1 - e^{-4\alpha t}\right)$$



Probability

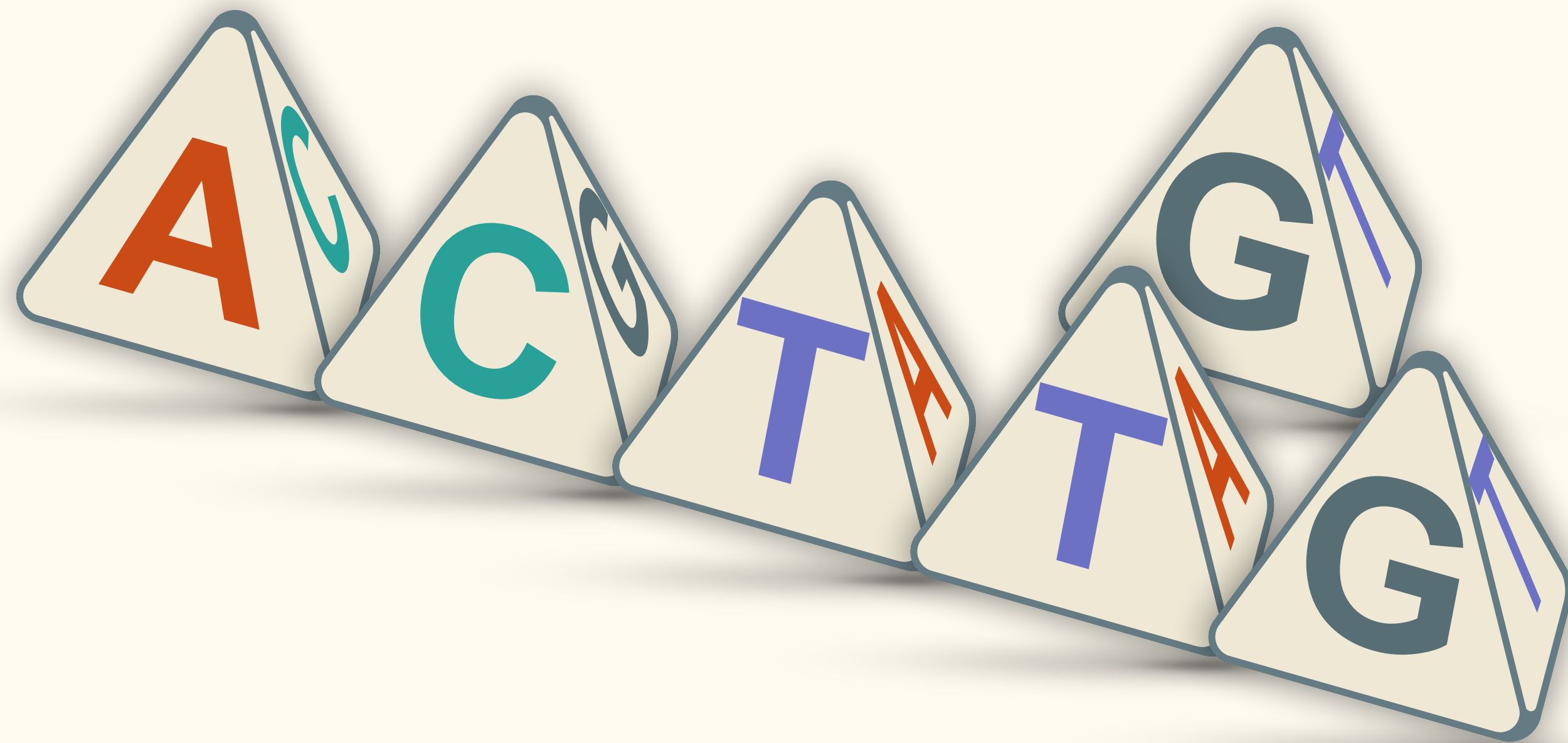
$$\times P(\text{T} \rightarrow \text{G} \mid \text{AC}, \text{Rate}, \text{Time})$$

The equation is annotated with three orange arrows pointing downwards to the parameters: 'Rate' points to the thermometer icon, and 'Time' points to the clock icon. The 'AC' in the equation is represented by a die with 'A' and 'C' on its faces.



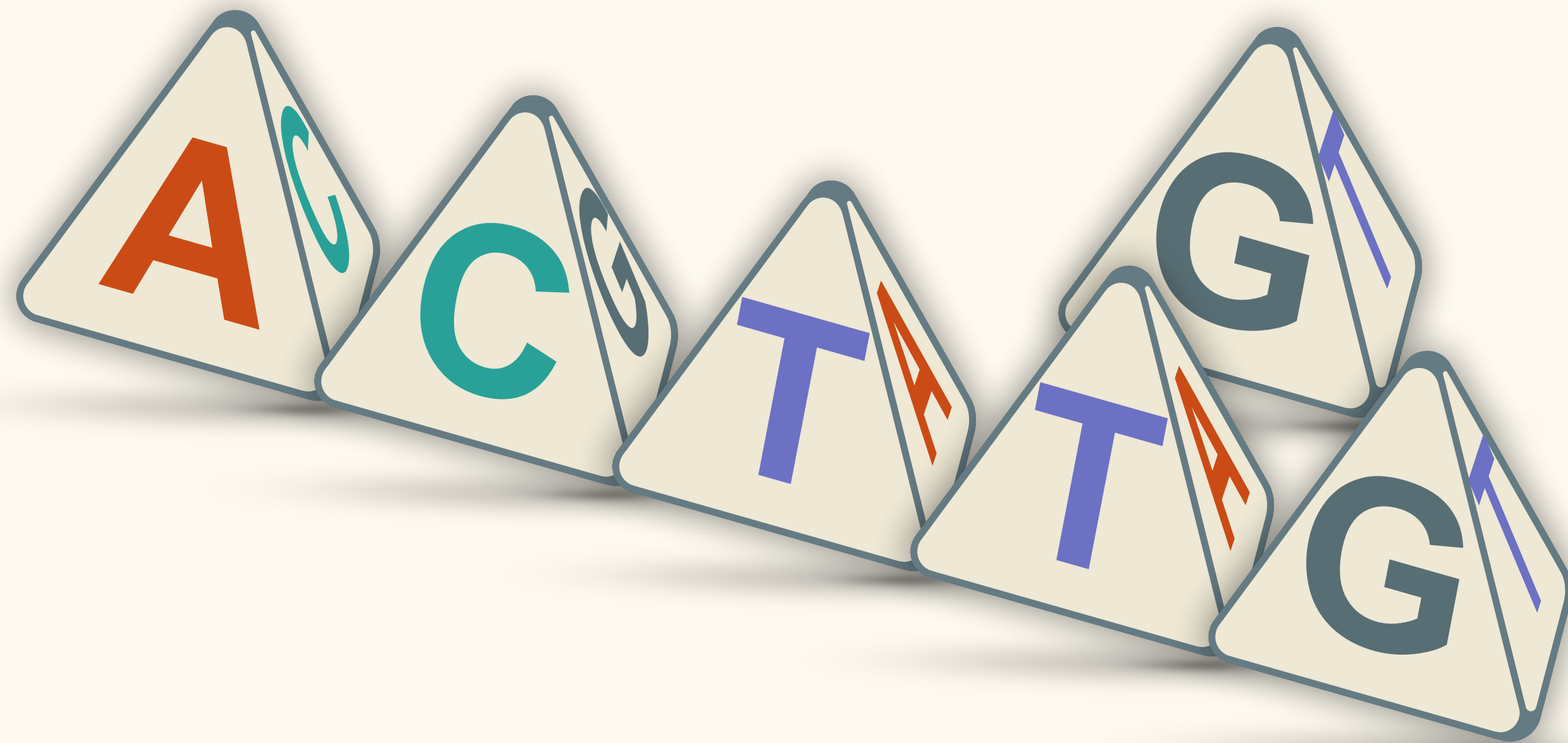
Probability

$$\times P(\mathbf{T} \rightarrow \mathbf{G} \mid \triangle \mathbf{A}, \text{Rate} = 1, \text{Time} = 1)$$



Probability

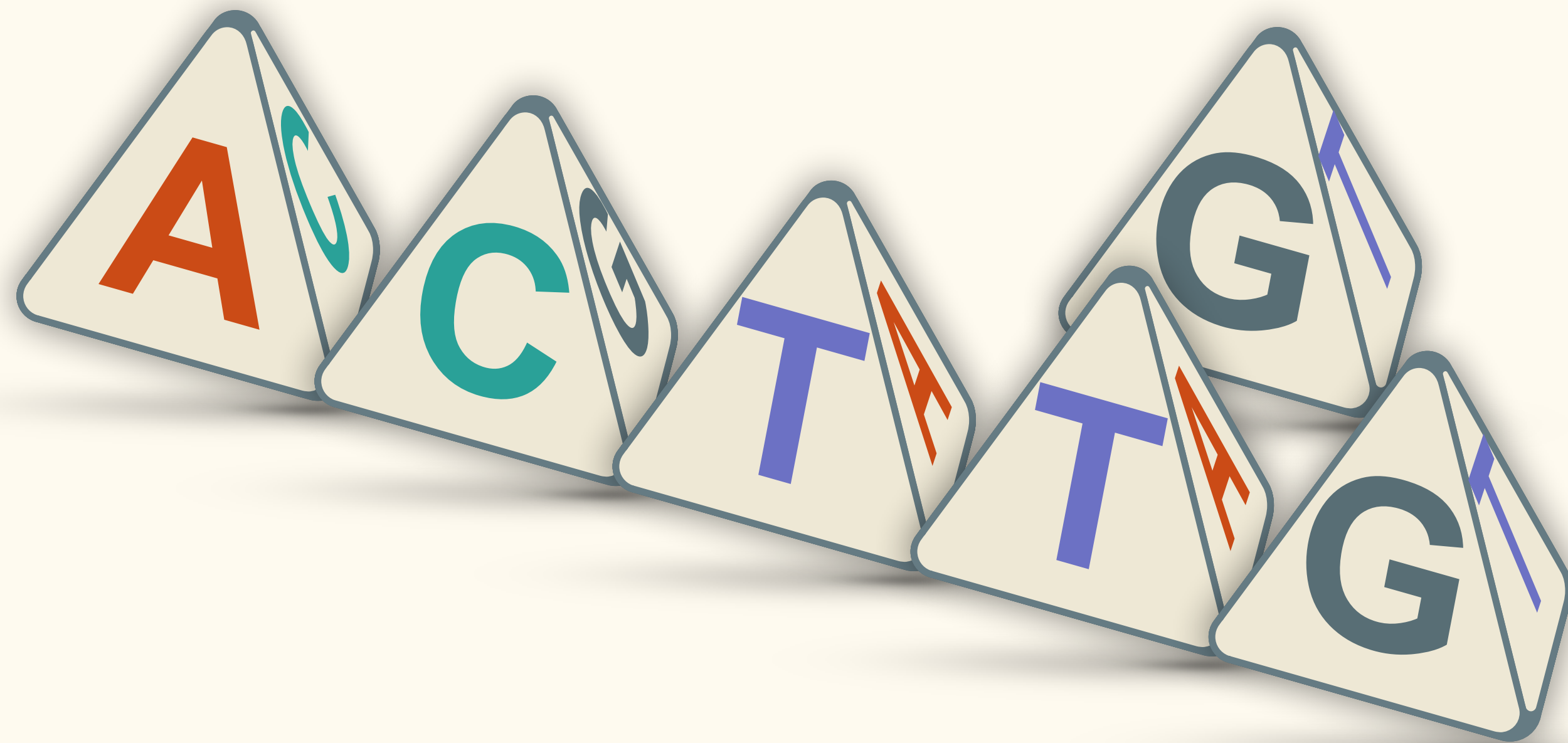
$$\times P(\mathbf{T} \rightarrow \mathbf{G} \mid \mathbf{M}_1)$$



Probability

Rate Time

$$\left(\frac{1}{4}\right) \times \left(1 - e^{-4\alpha t}\right)$$



Probability



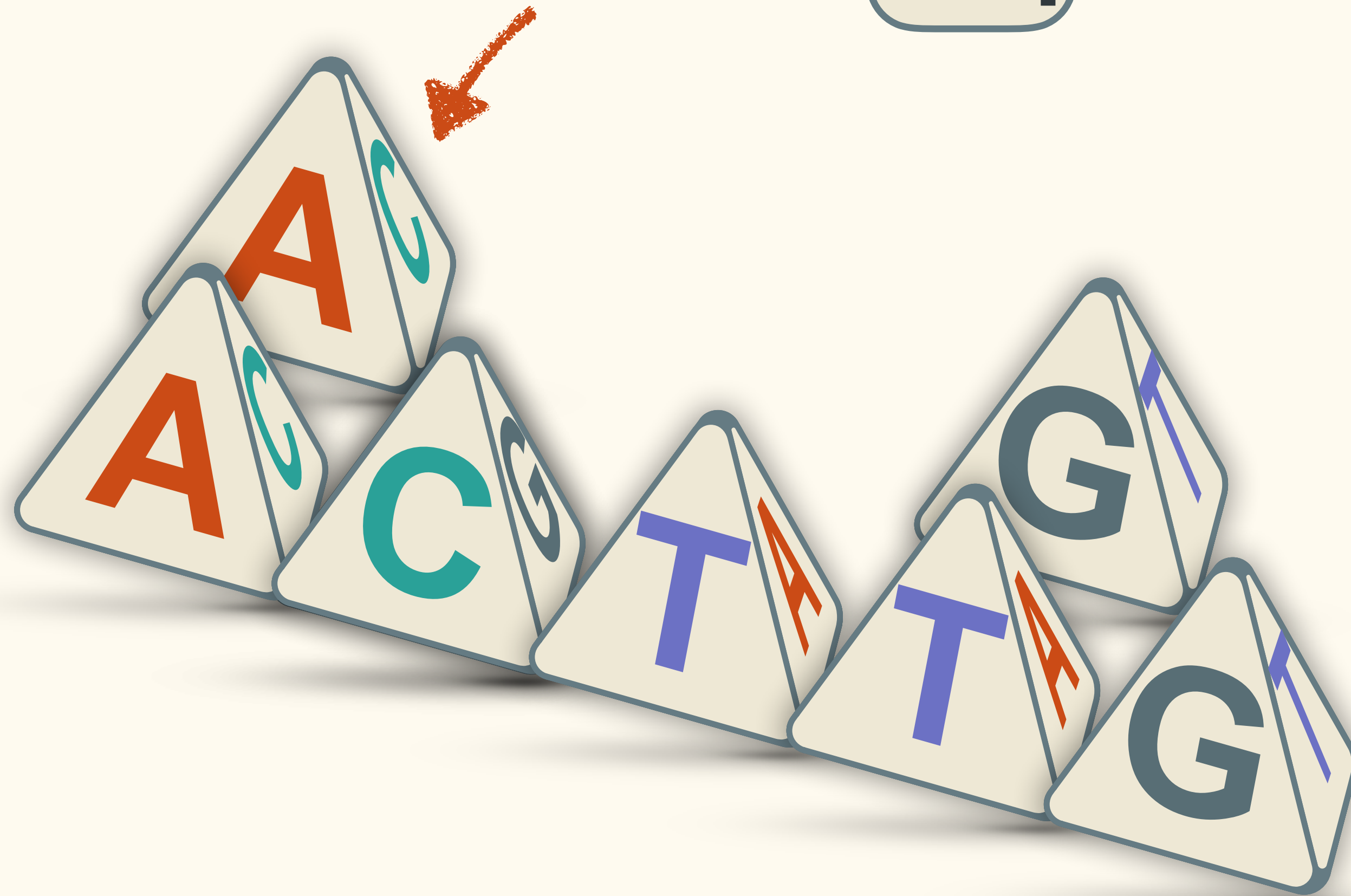
Probability

$$P\left(\begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \mid M_1\right) = ?$$



Probability

$$\times P(\mathbf{A} \rightarrow \mathbf{A} \mid \mathbf{M}_1)$$



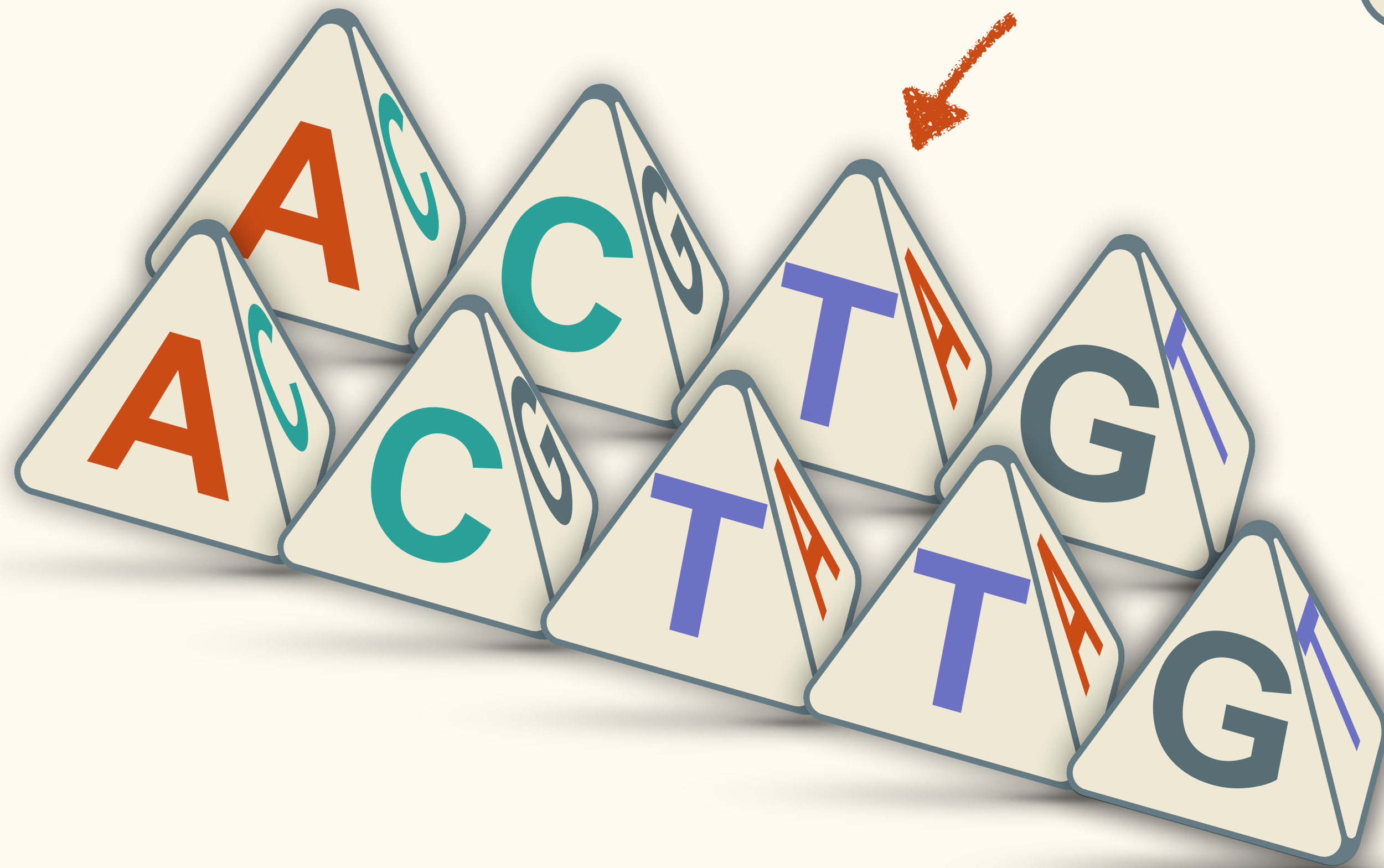
Probability

$$\times P(\text{C} \rightarrow \text{C} \mid \boxed{M_1})$$

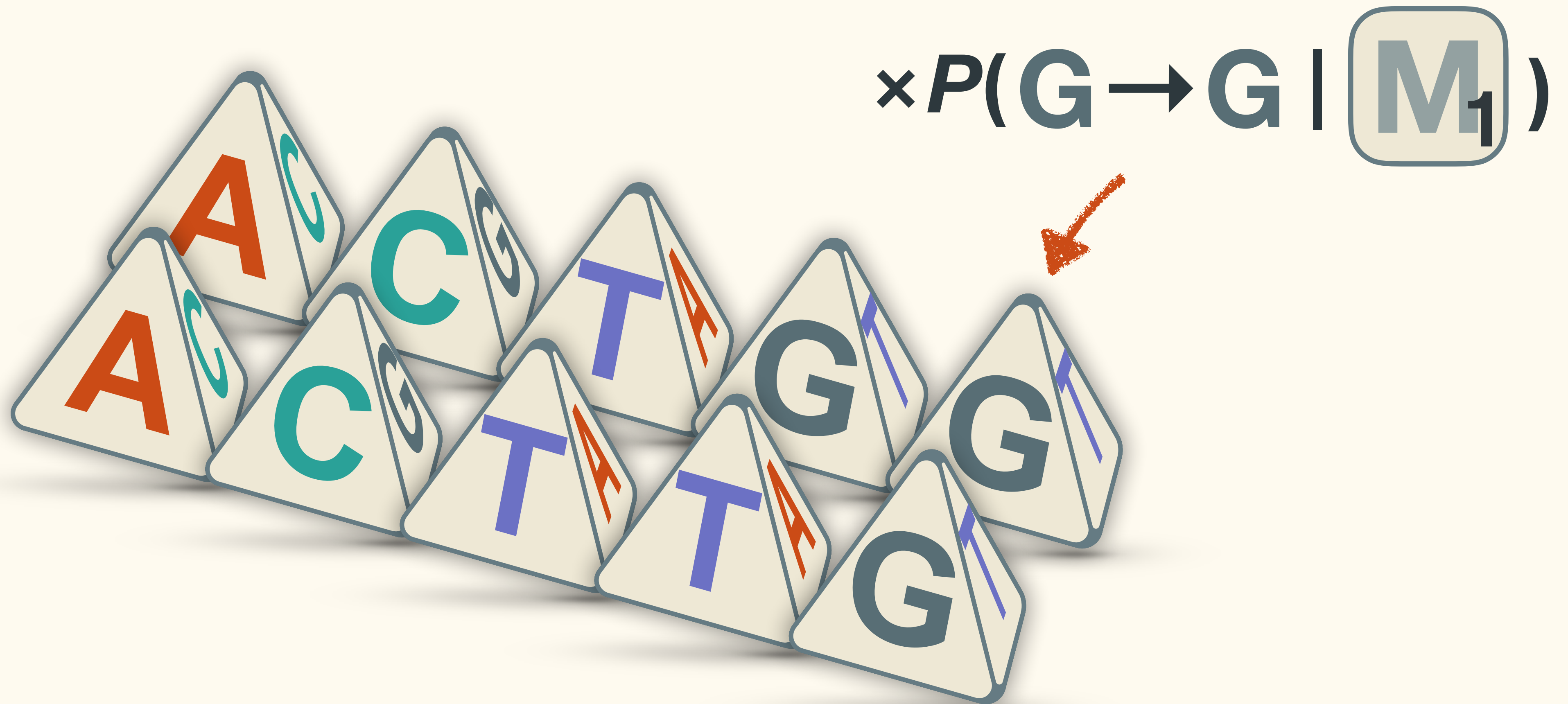


Probability

$$\times P(T \rightarrow T \mid M_1)$$



Probability



Probability

$$\times P(A \rightarrow A \mid M_1)$$



Probability

Rate

$$\left(\frac{1}{4}\right) \times (1 - e^{-4\alpha t}) + e^{-4\alpha t}$$



Probability

Time

$$\left(\frac{1}{4}\right) \times (1 - e^{-4\alpha t}) + e^{-4\alpha t}$$



Probability

0.2637

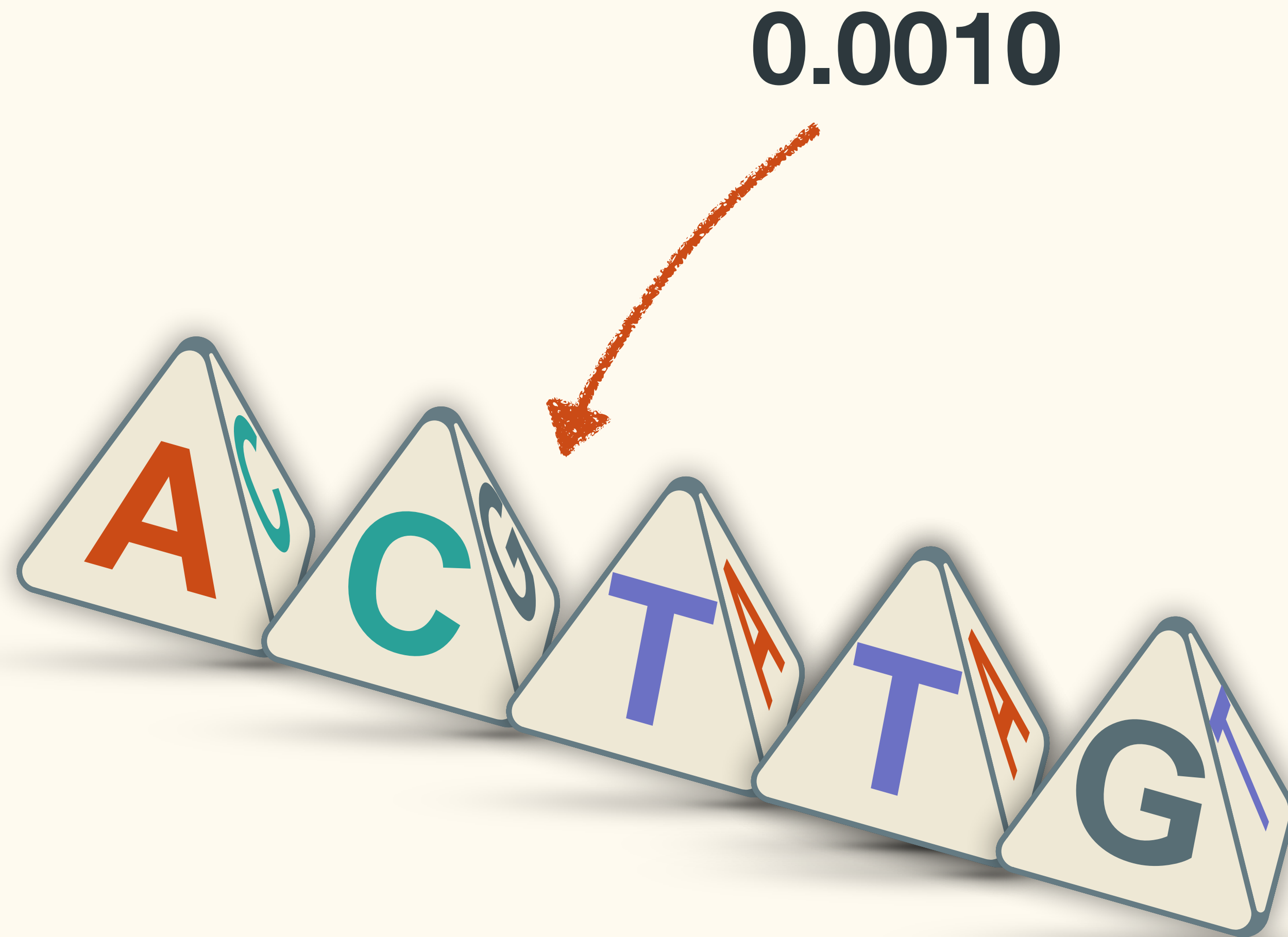


Probability

$$P\left(\begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \mid M_1\right) = ?$$

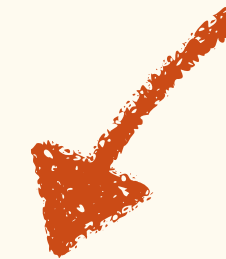


Probability



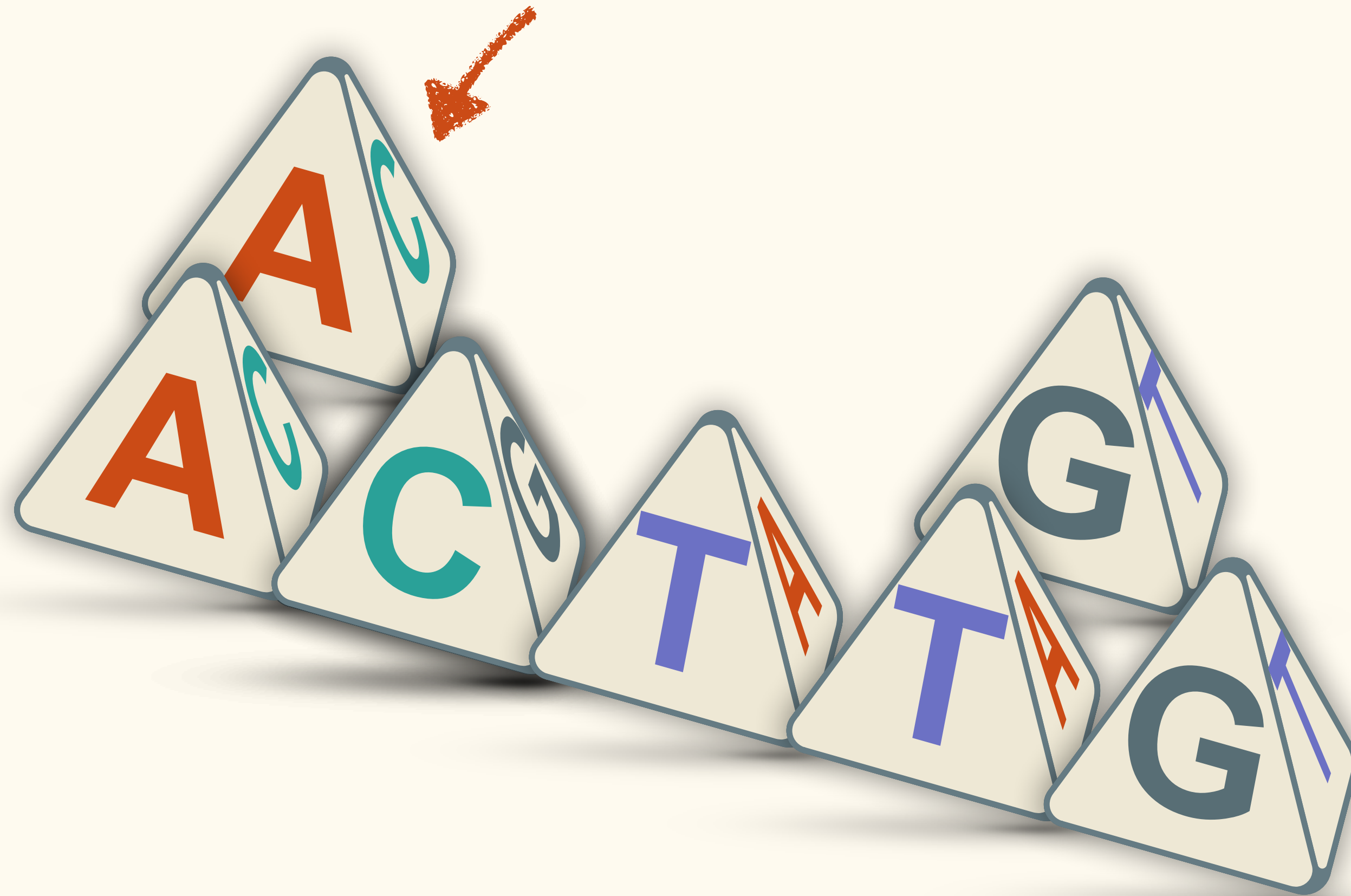
Probability

$\times 0.2454$



Probability

$\times 0.2637$

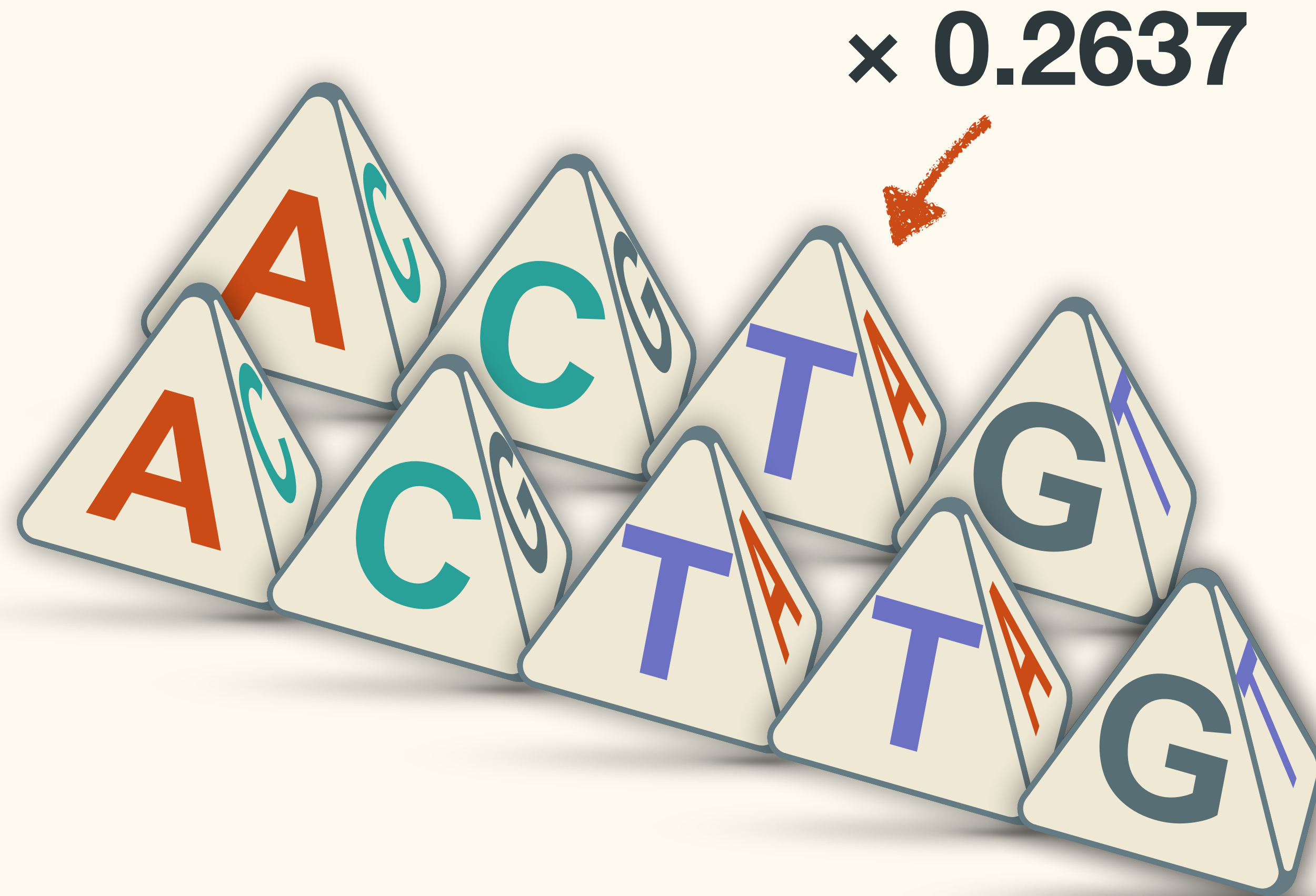


Probability

$\times 0.2637$



Probability

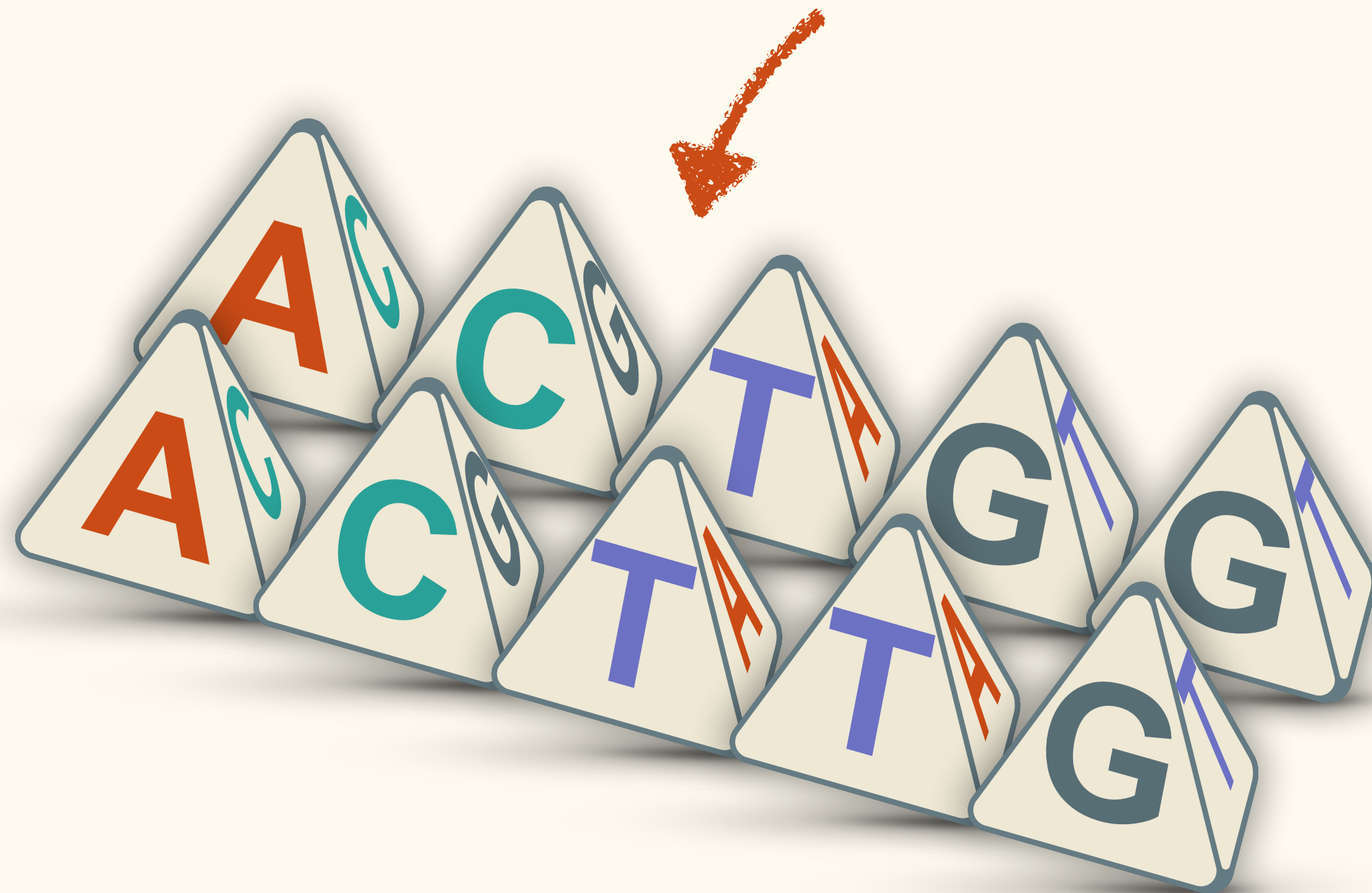


Probability



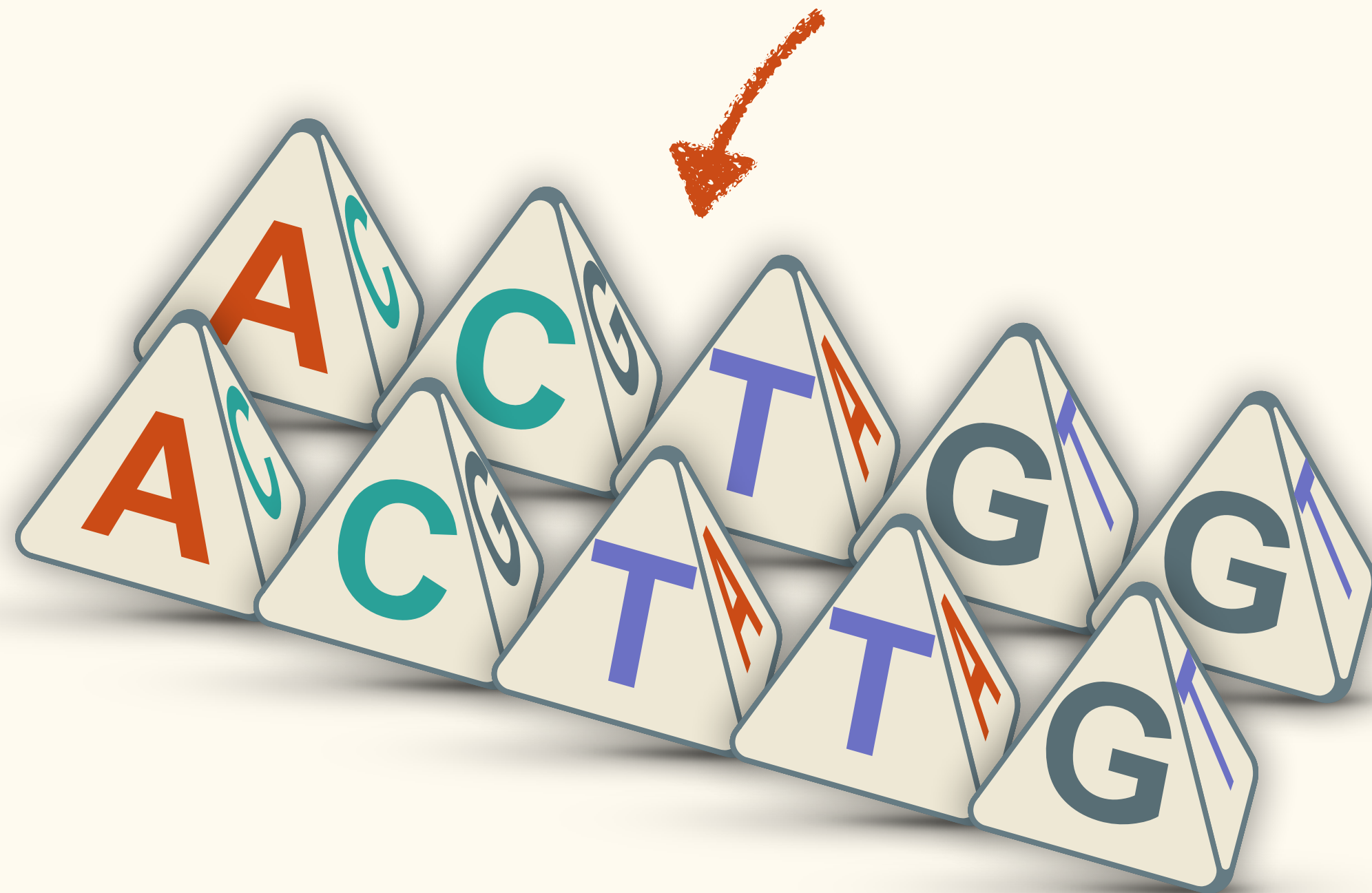
Probability

$$P\left(\begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \mid M_1\right) = 0.0000015$$



Likelihood

$$L(\boxed{M_1} \mid \begin{matrix} \text{ACTTTG} \\ \text{ACTGGG} \end{matrix}) = 0.0000015$$



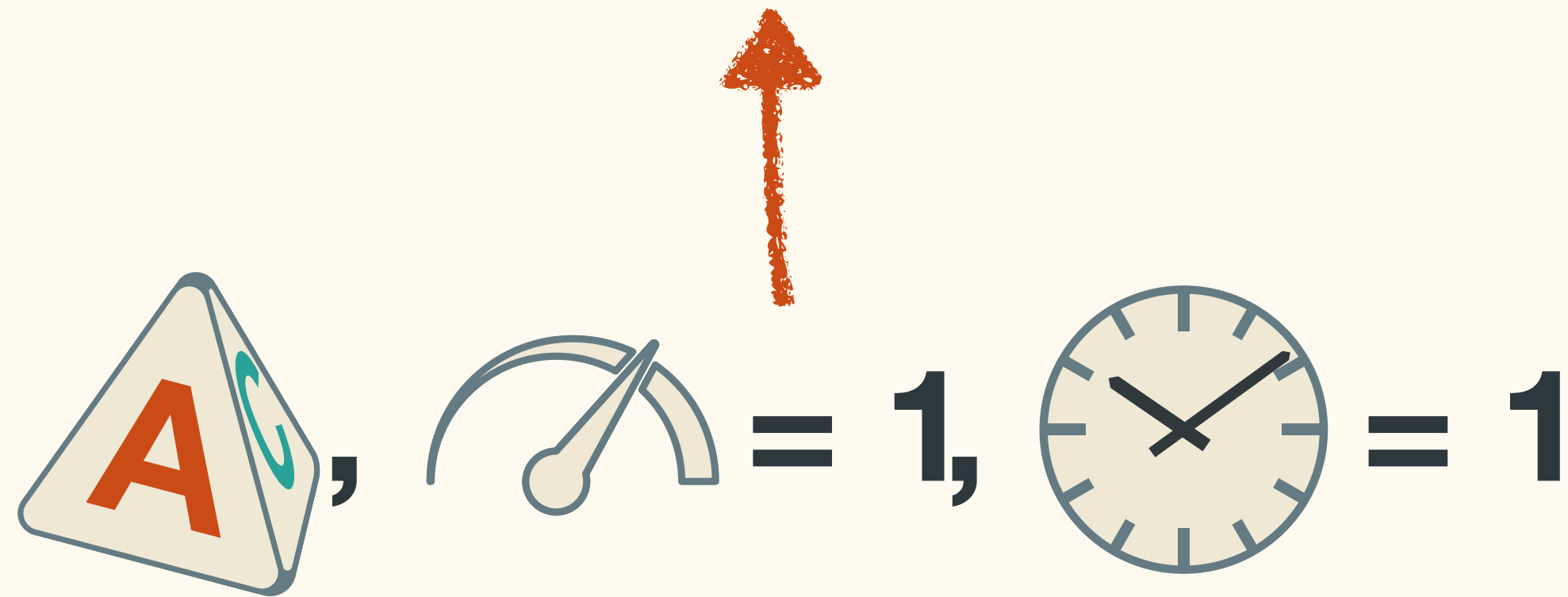
Likelihood

$$\log \left(L \left(\boxed{M_1} \mid \begin{array}{l} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$




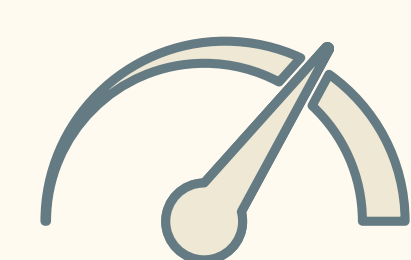
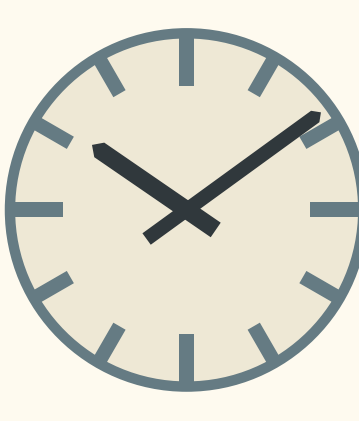
Likelihood

$$\log \left(L \left(\boxed{M_1} \mid \begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$




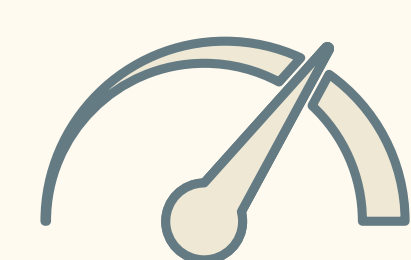
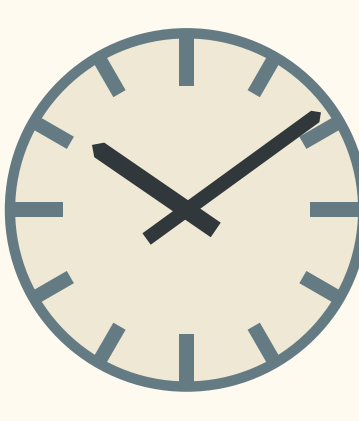
Likelihood

$$\log \left(L \left(\boxed{M_2} \mid \begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$

 ,  = 0.1,  = 1

Likelihood

$$\log \left(L \left(\boxed{M_2} \mid \begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -10.3$$

 ,  = 0.1,  = 1

Likelihood

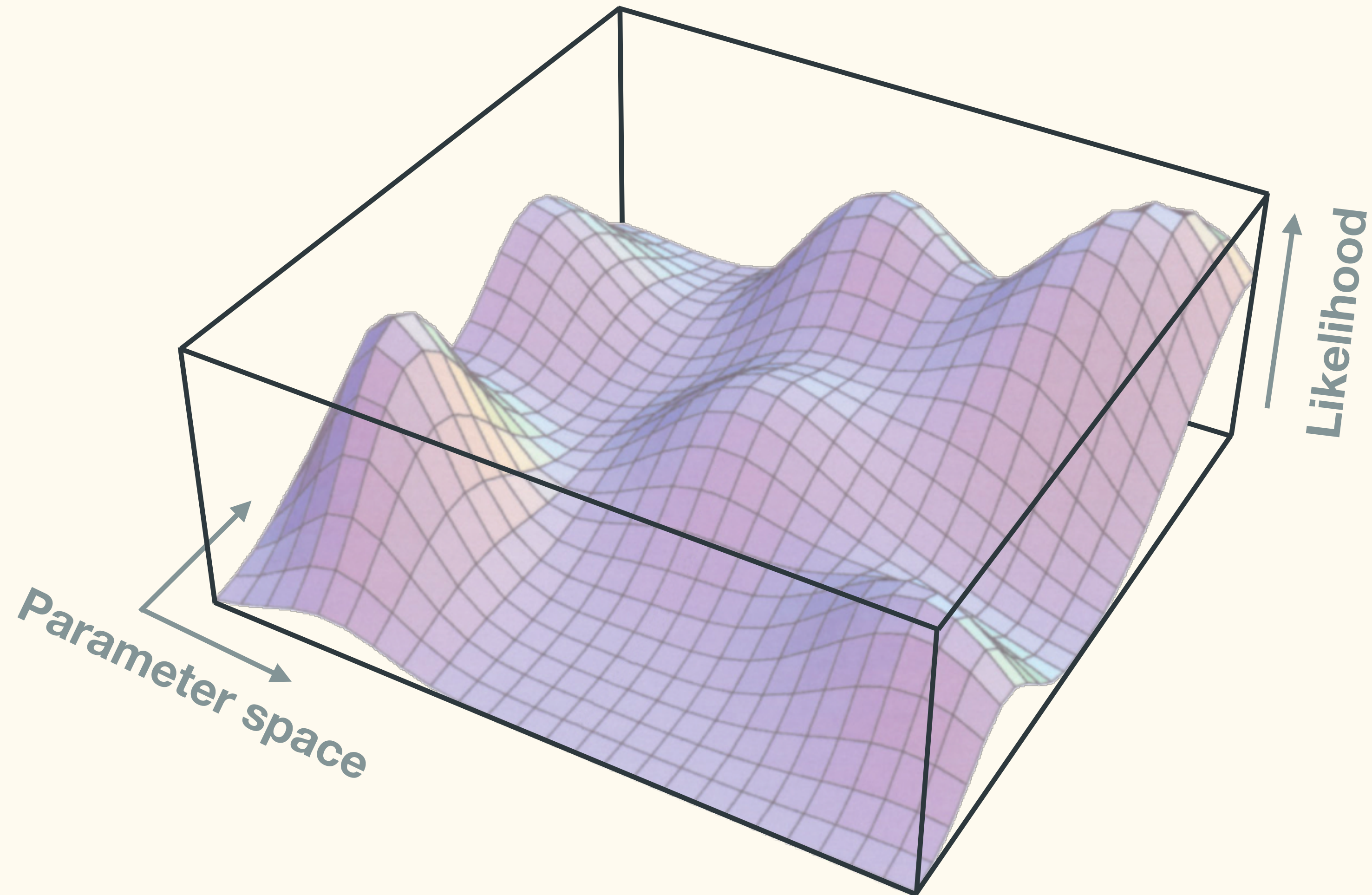
$$\log \left(L \left(\boxed{M_1} \mid \begin{array}{l} \text{ACTTTG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$

$$\log \left(L \left(\boxed{M_2} \mid \begin{array}{l} \text{ACTTTG} \\ \text{ACTGGG} \end{array} \right) \right) = -10.3$$

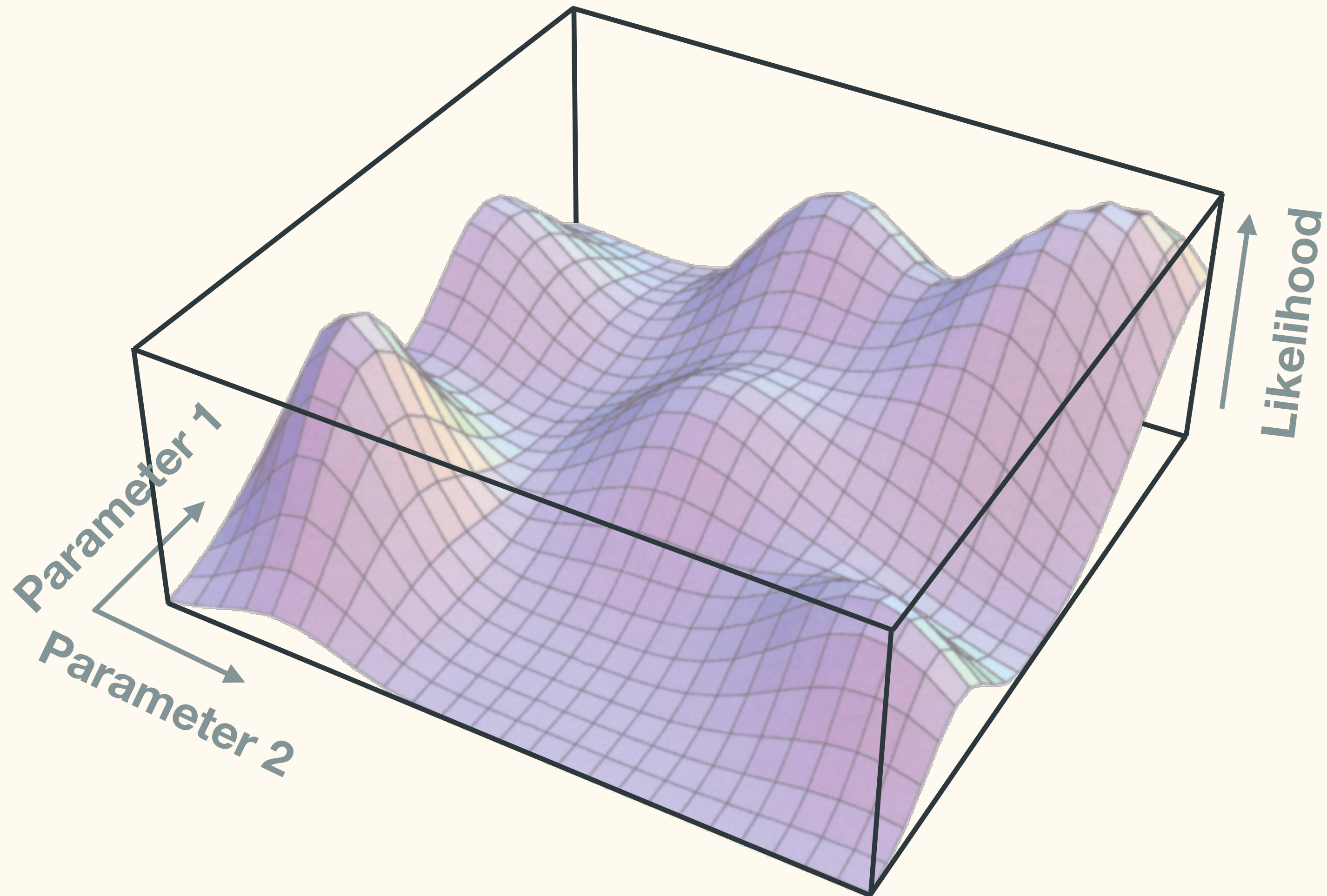
Estimating model parameters using likelihood

Maximum likelihood

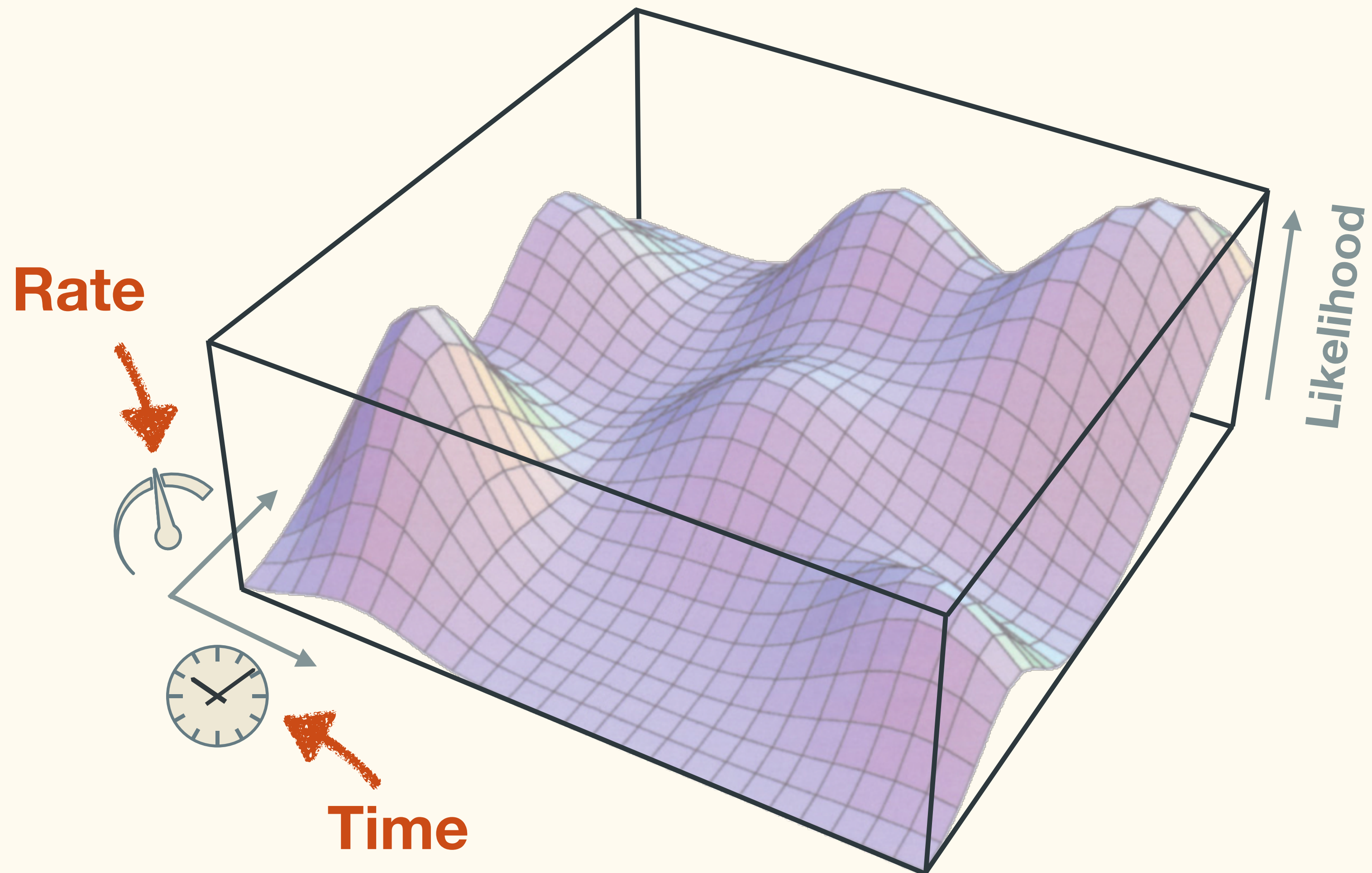
Likelihood surface



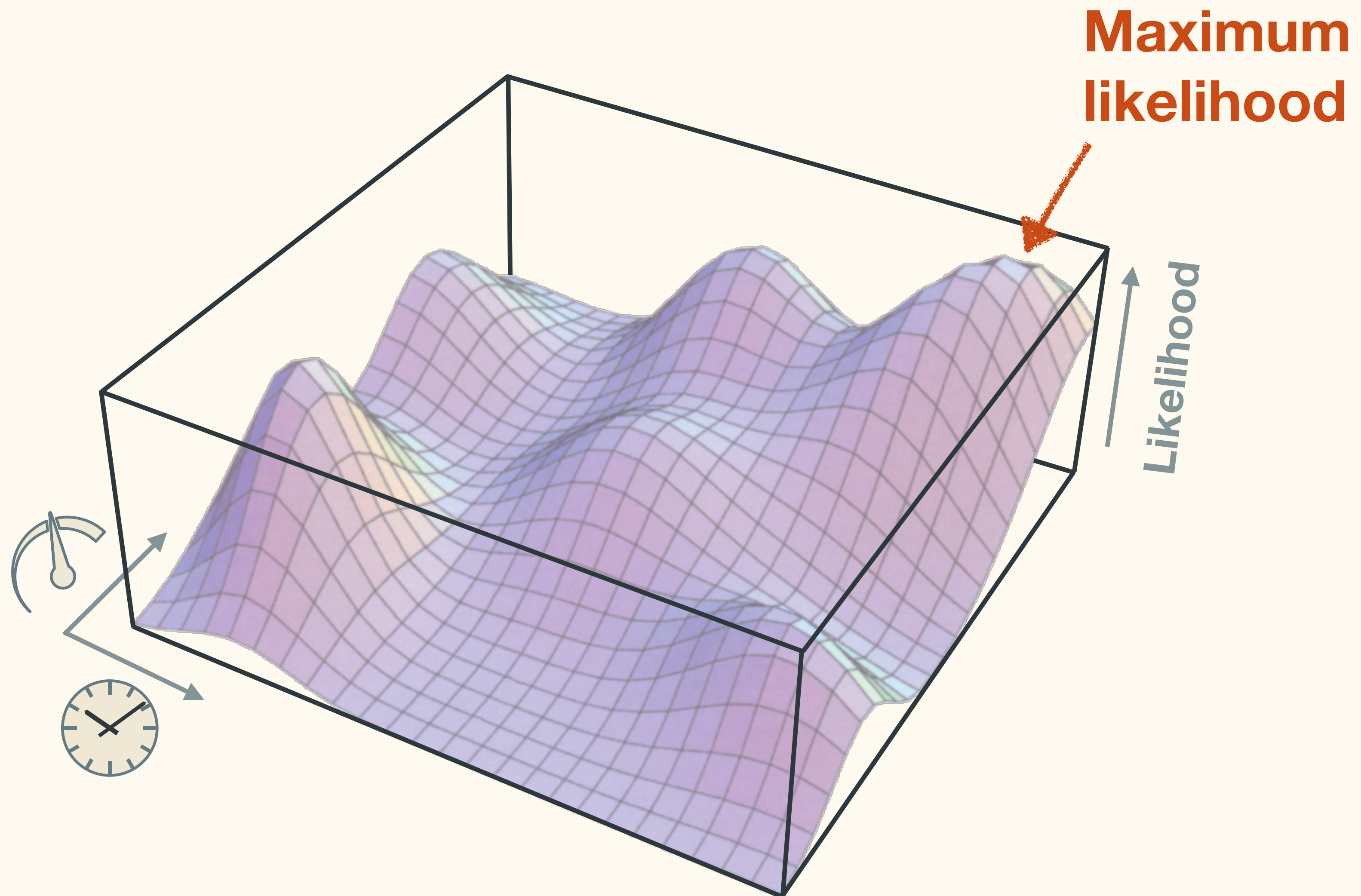
Likelihood surface



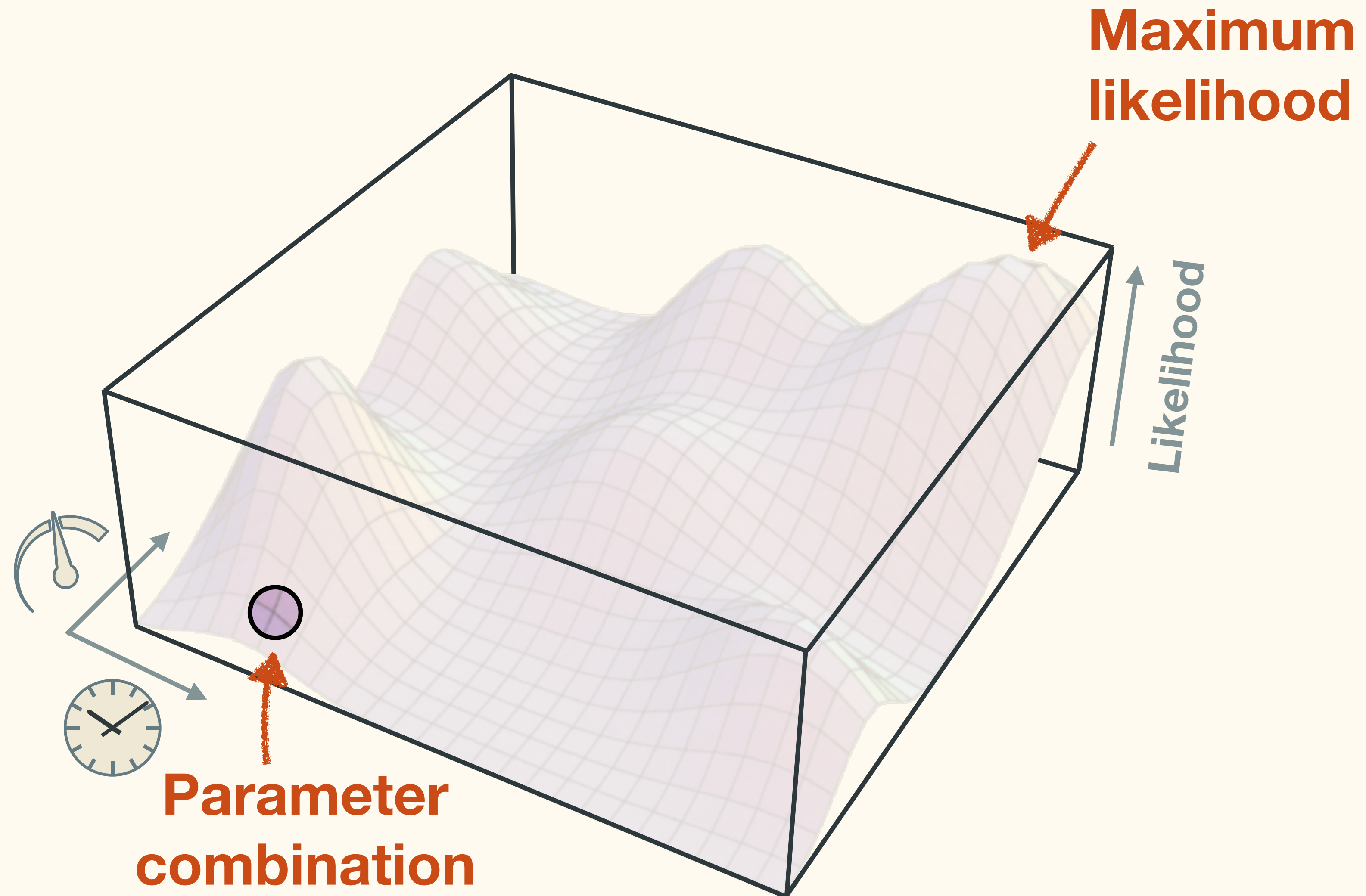
Likelihood surface



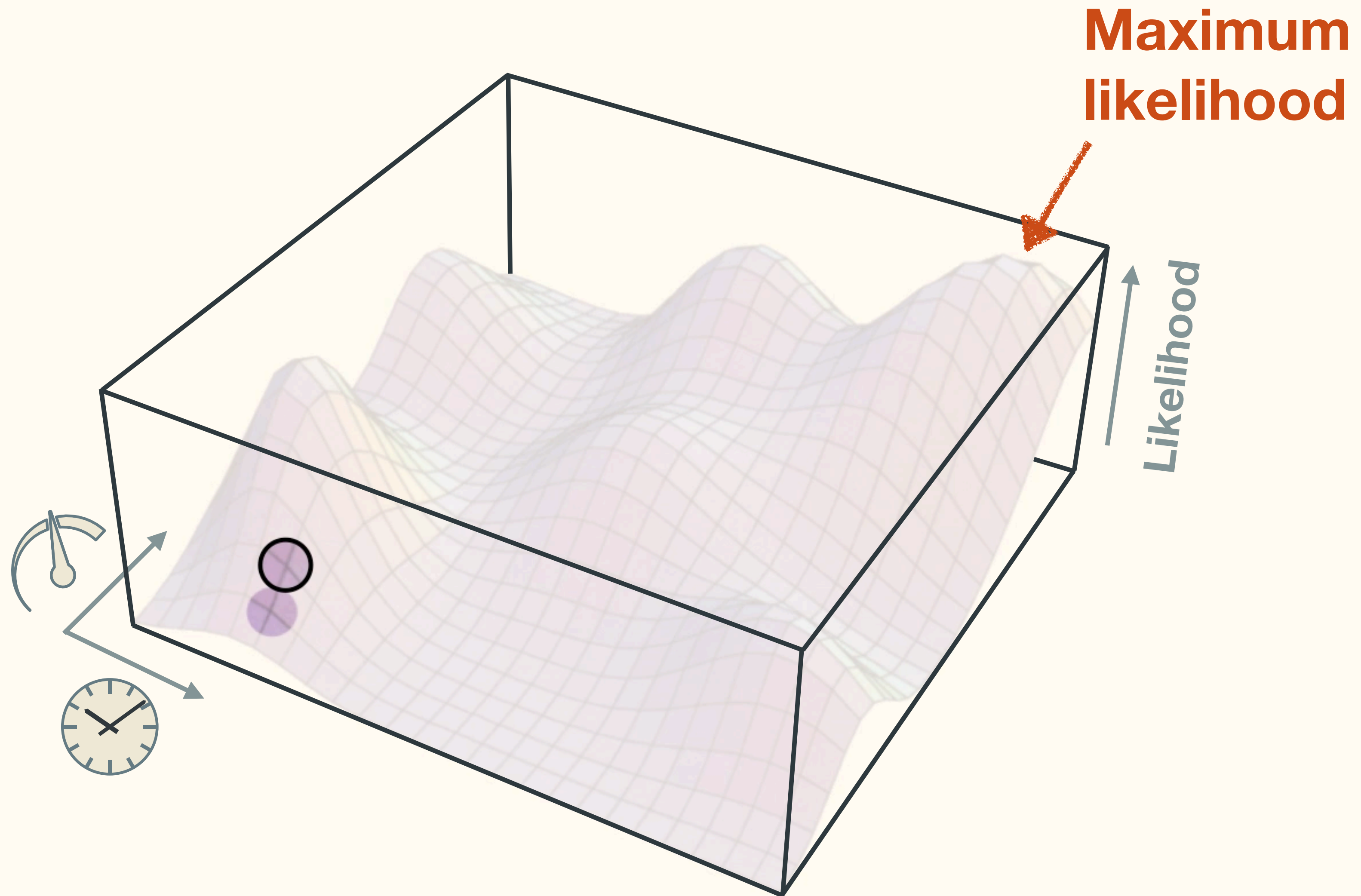
Likelihood surface



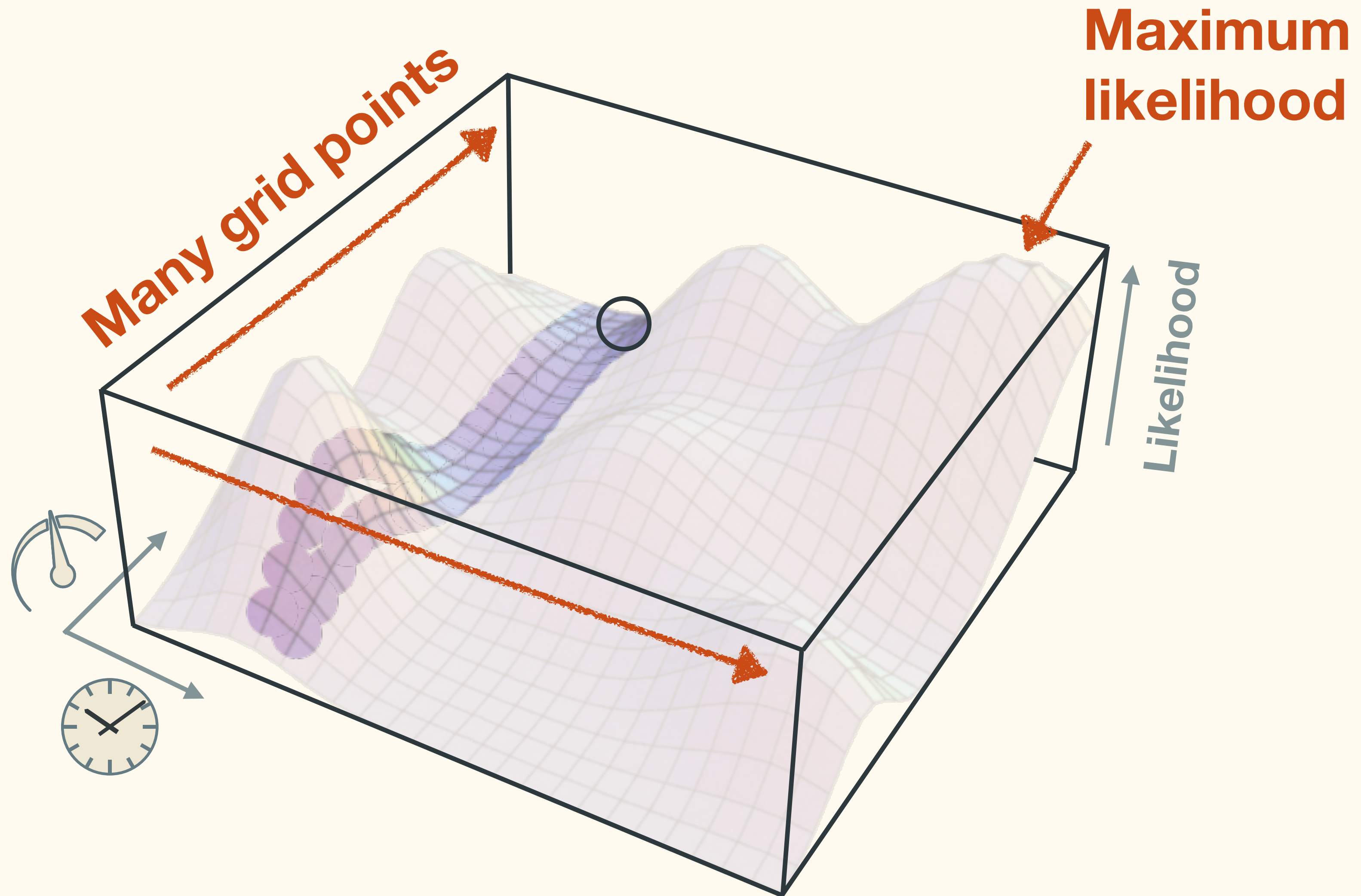
Likelihood surface



Grid search

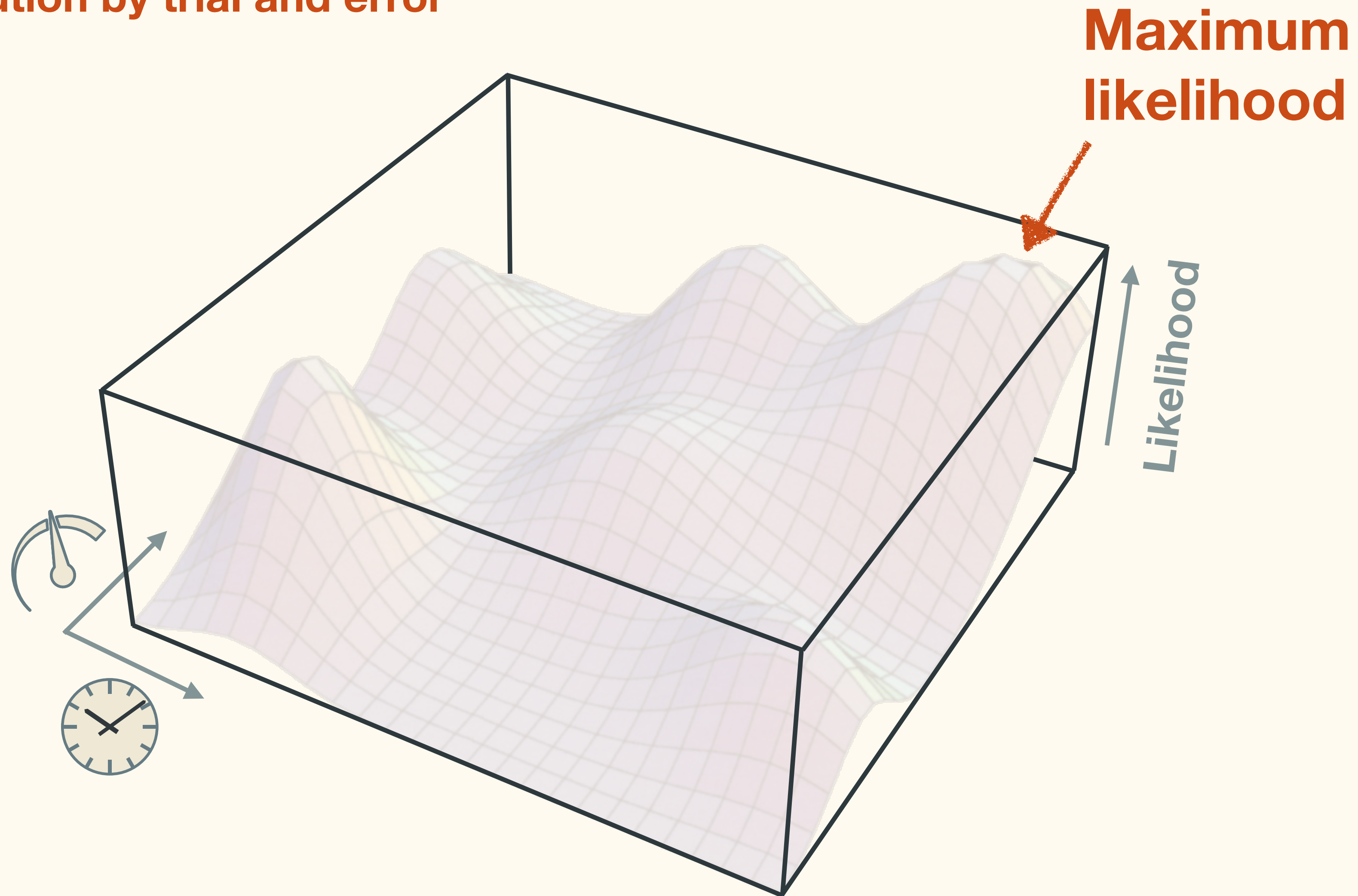


Grid search



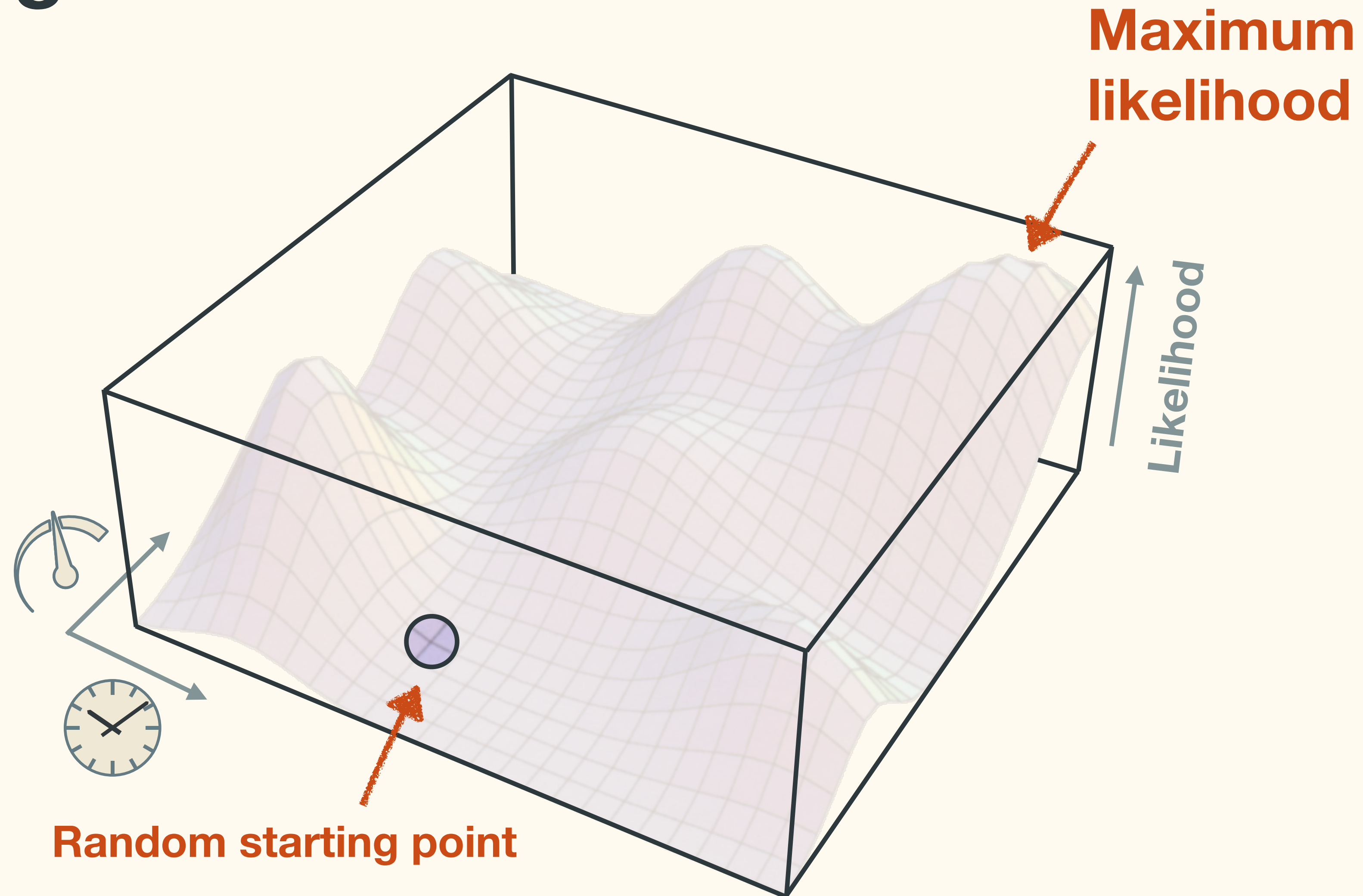
Heuristic search

“proceeding to a solution by trial and error”



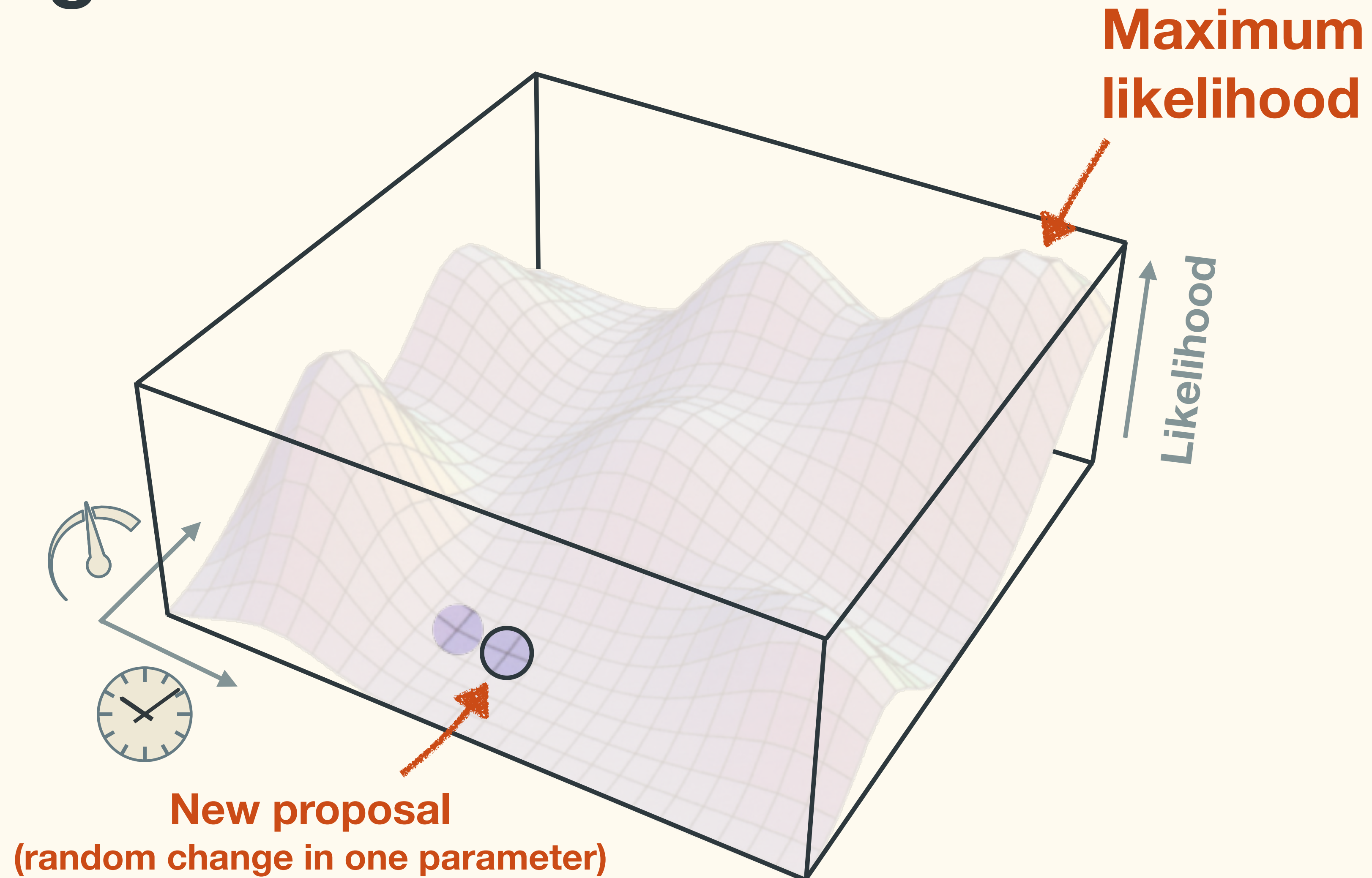
Heuristic search

Hill climbing



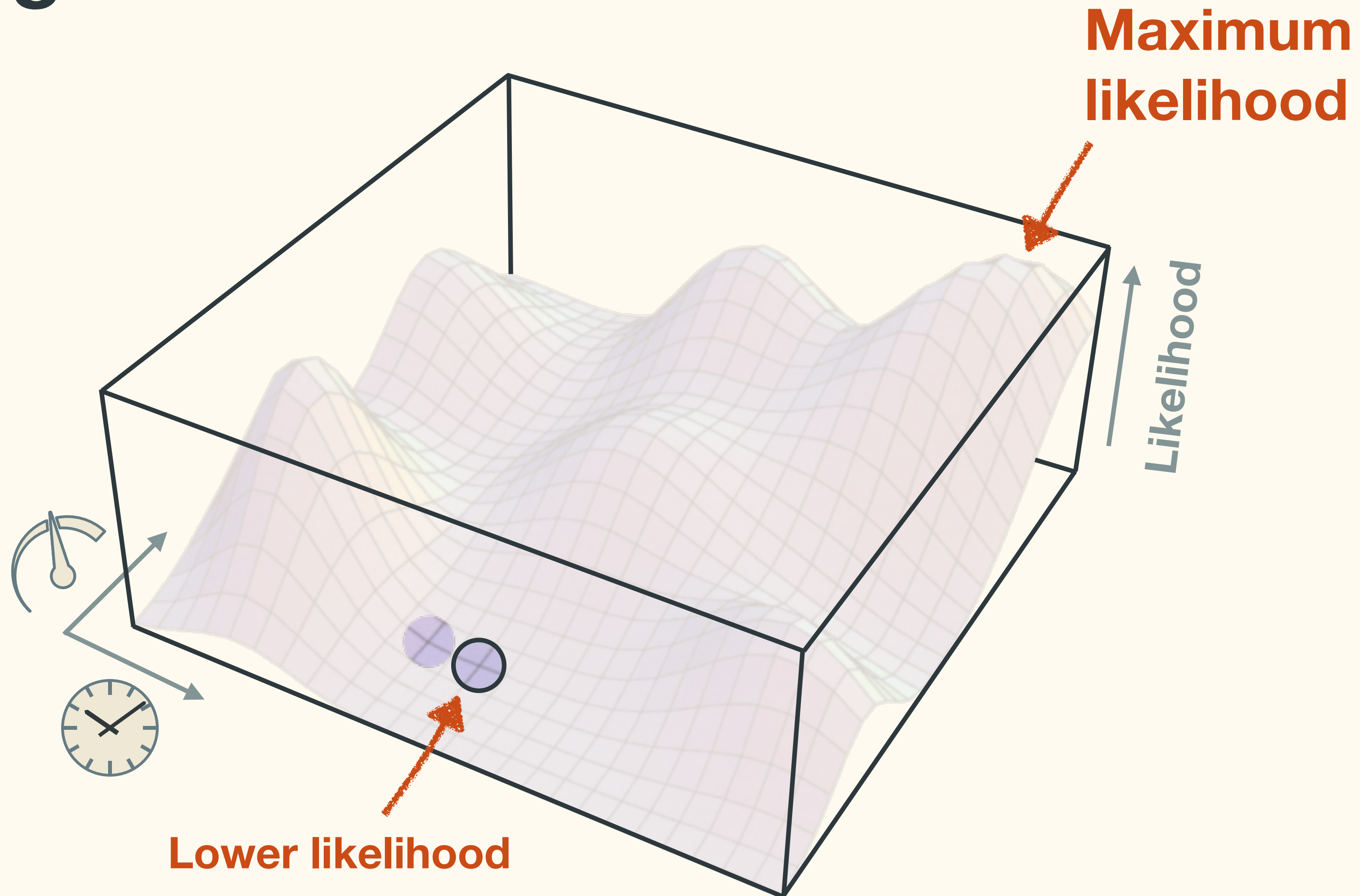
Heuristic search

Hill climbing



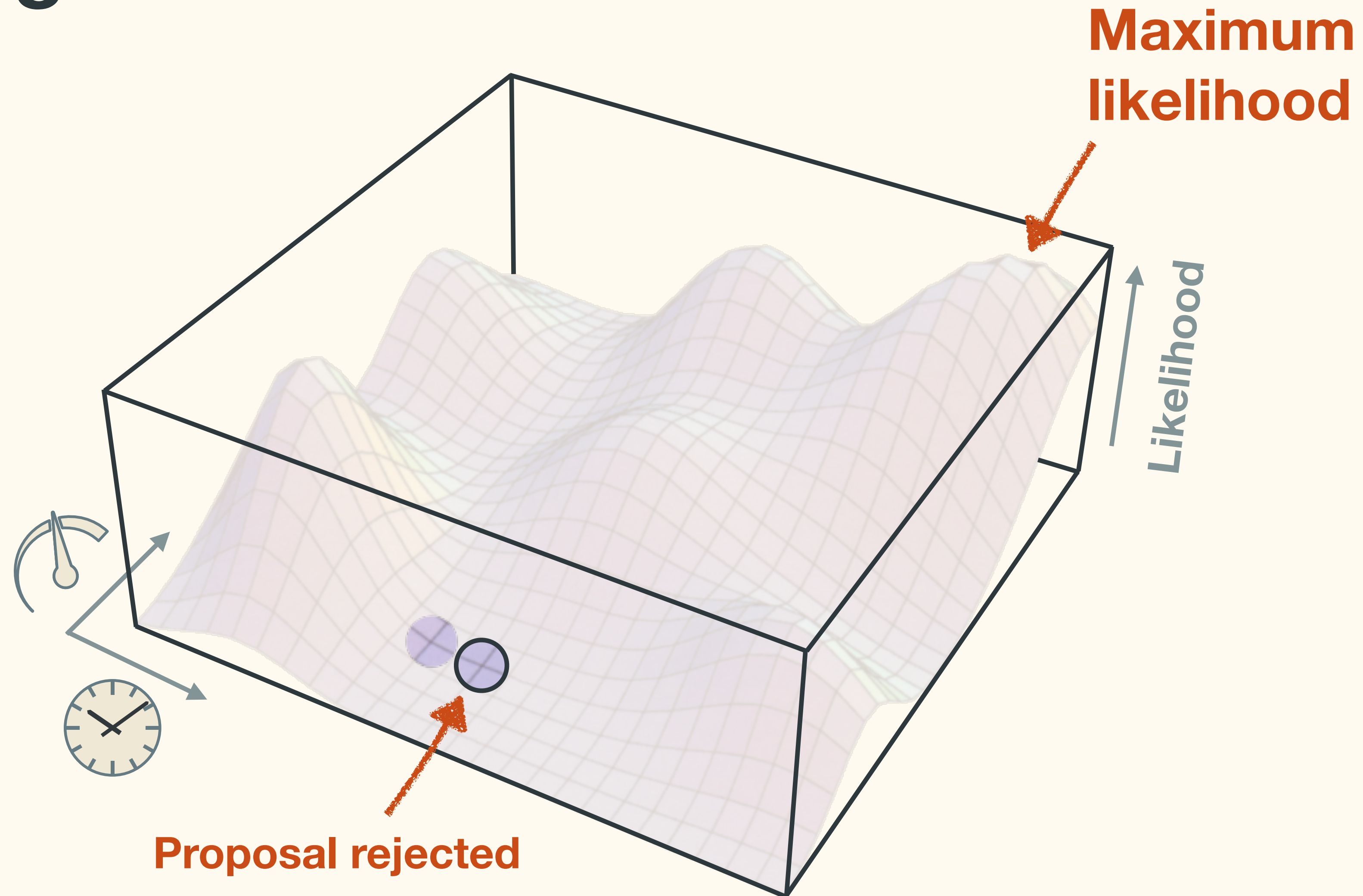
Heuristic search

Hill climbing



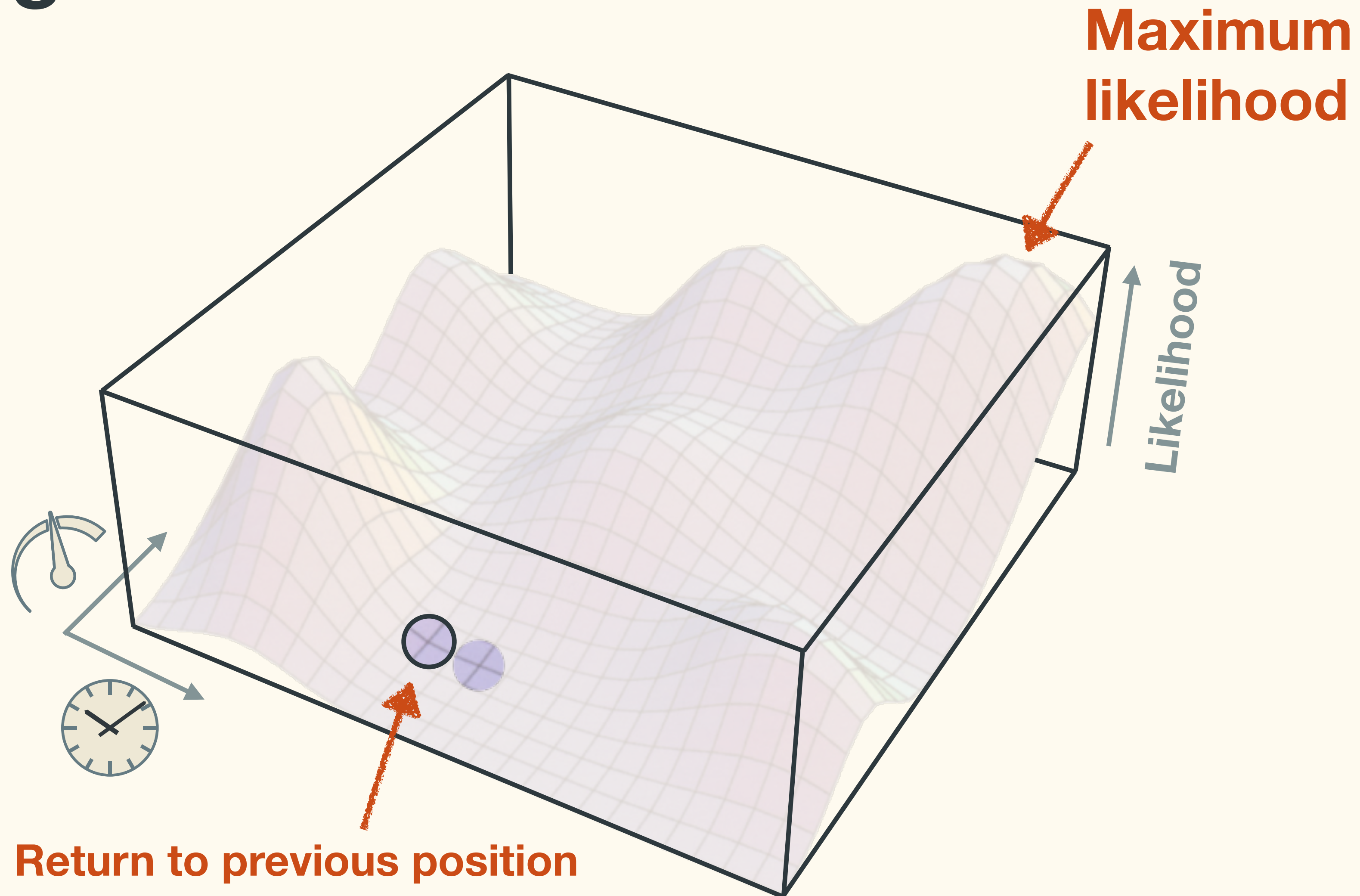
Heuristic search

Hill climbing



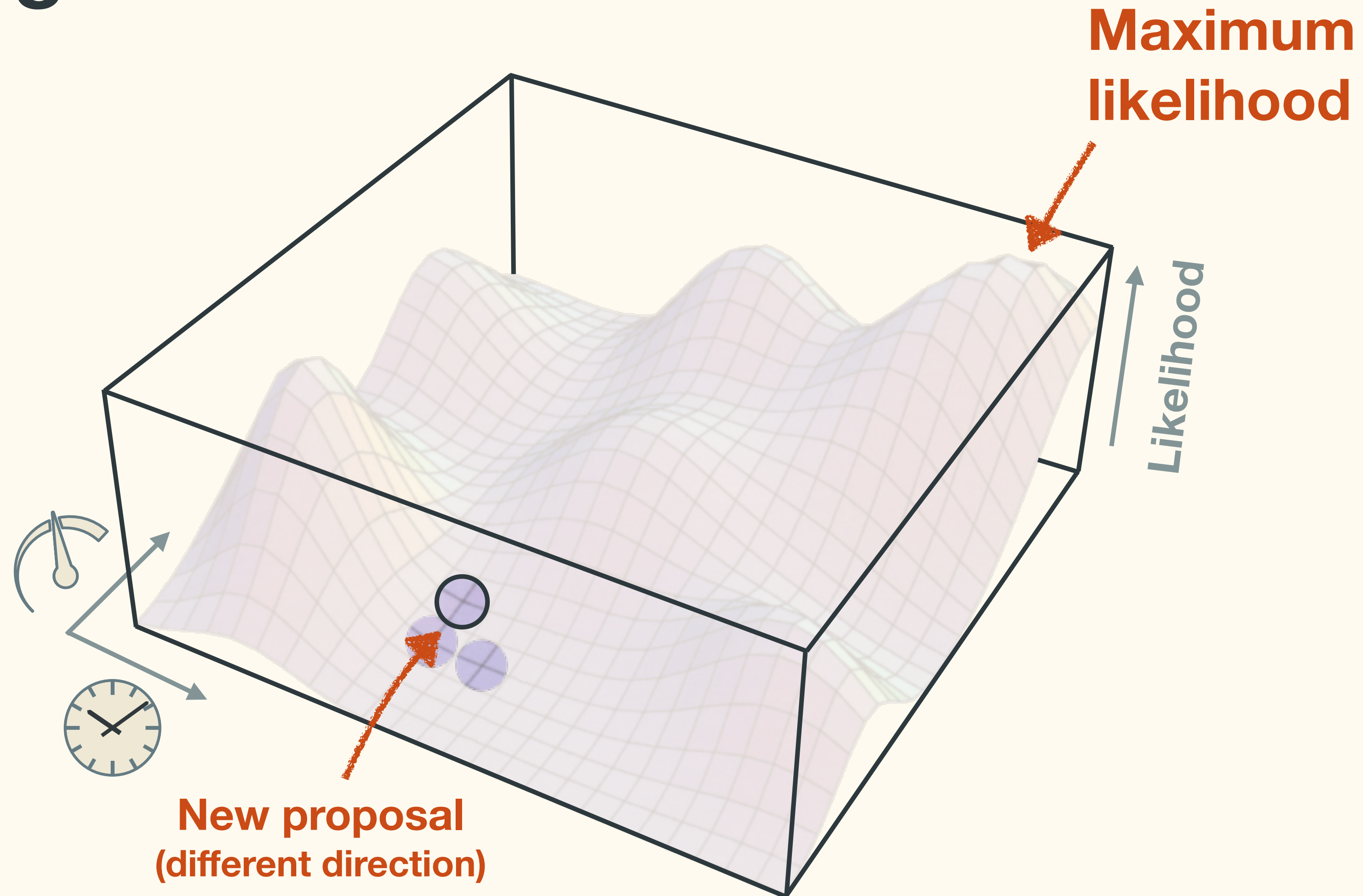
Heuristic search

Hill climbing



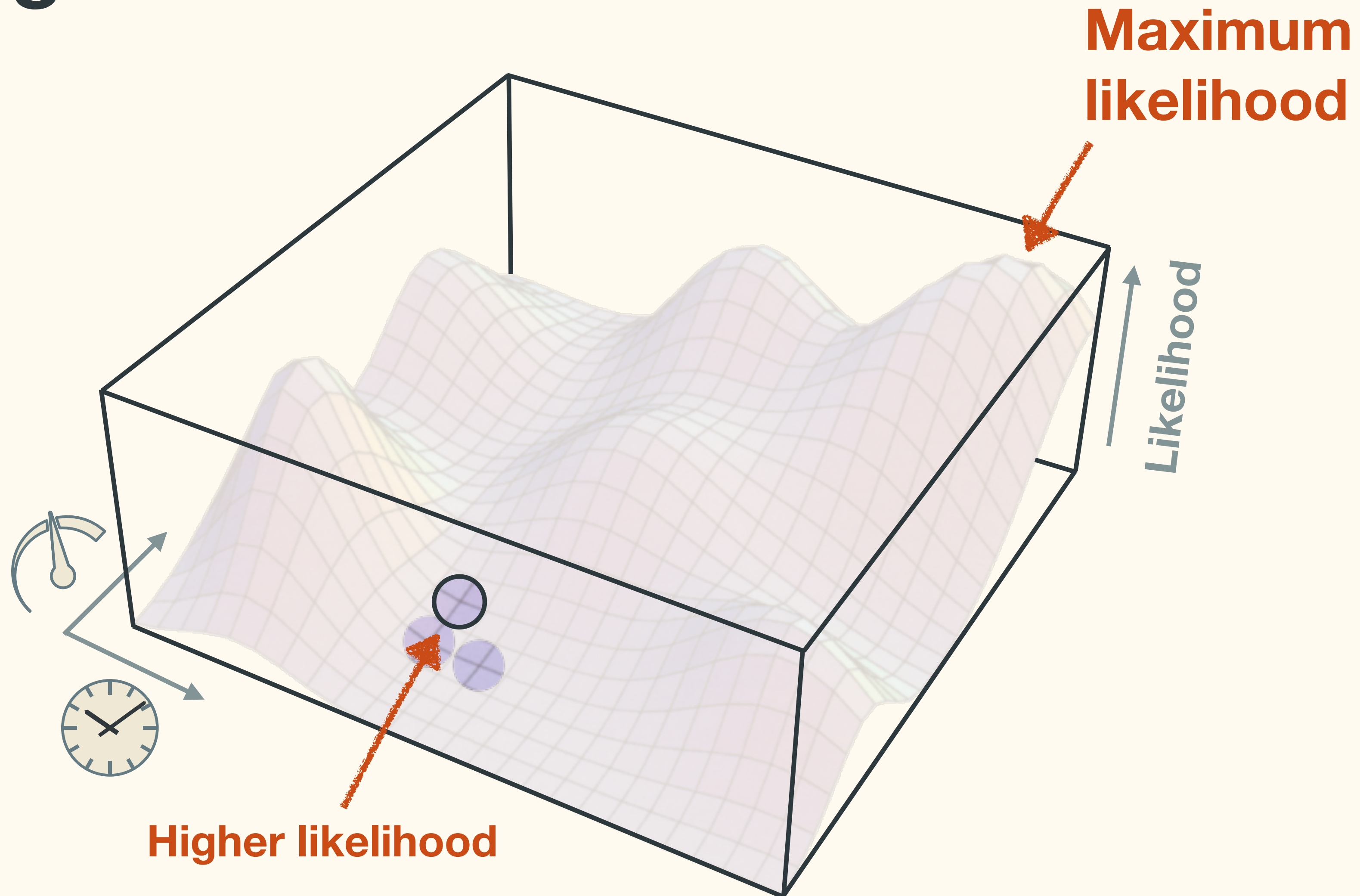
Heuristic search

Hill climbing



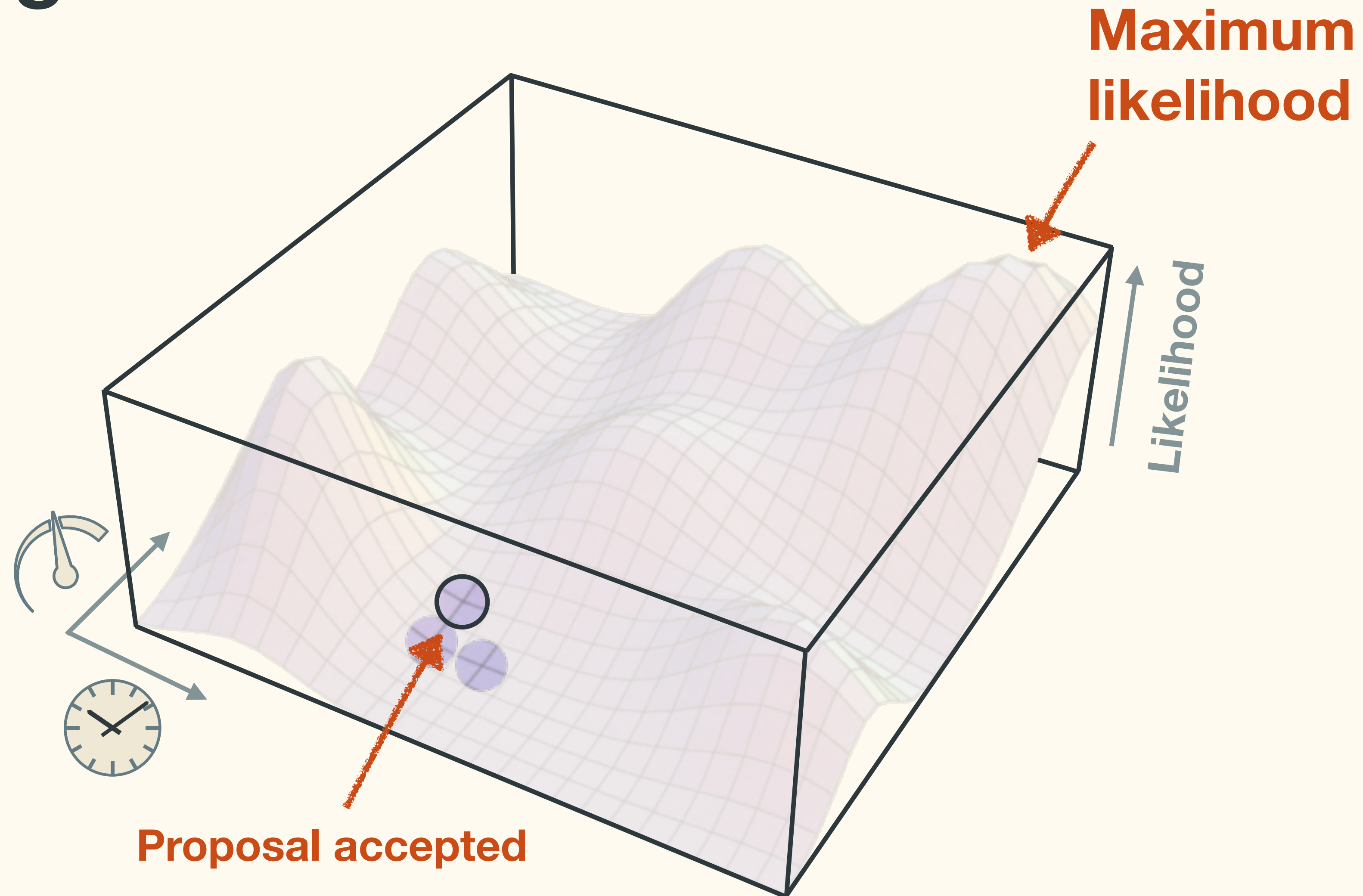
Heuristic search

Hill climbing



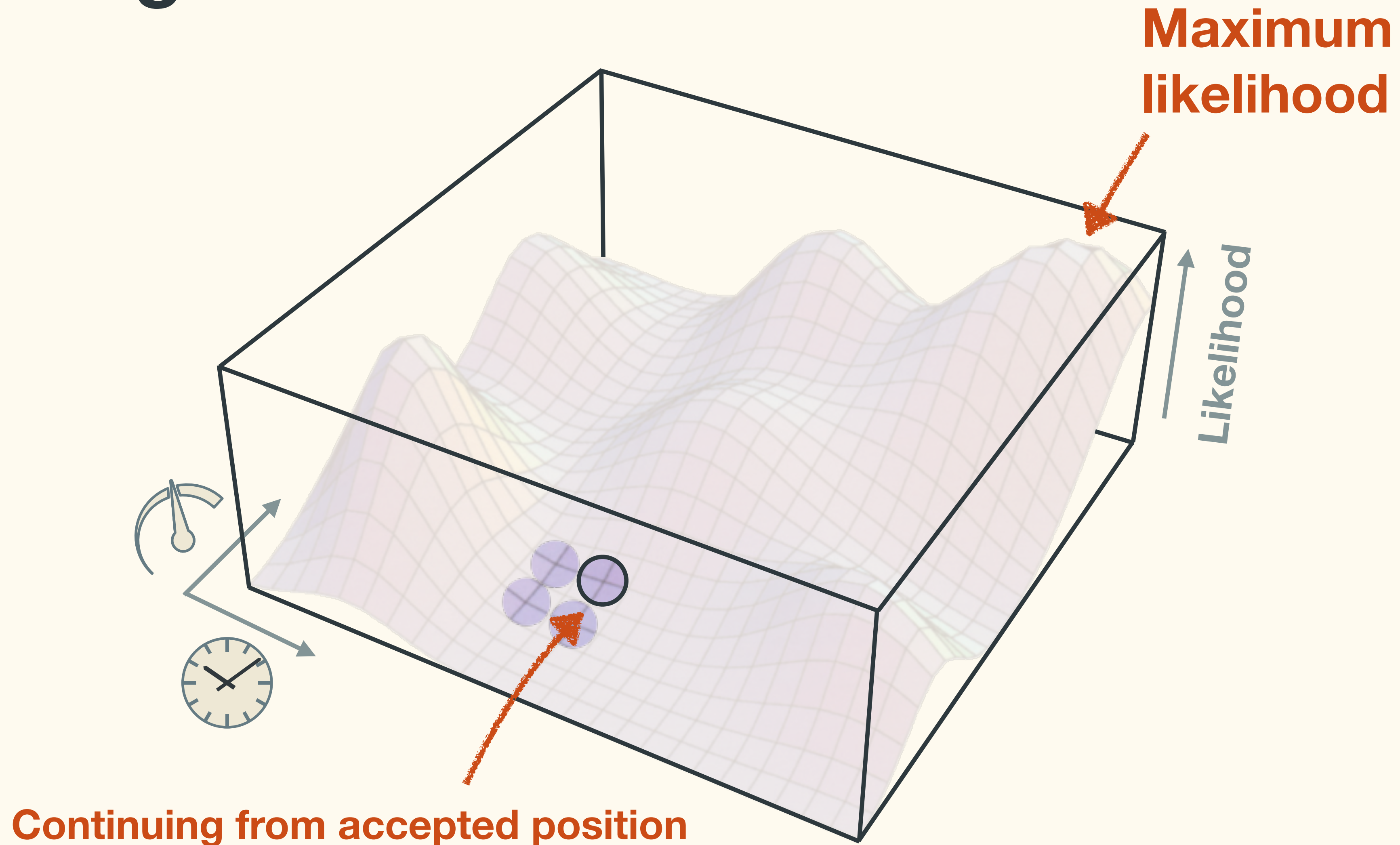
Heuristic search

Hill climbing



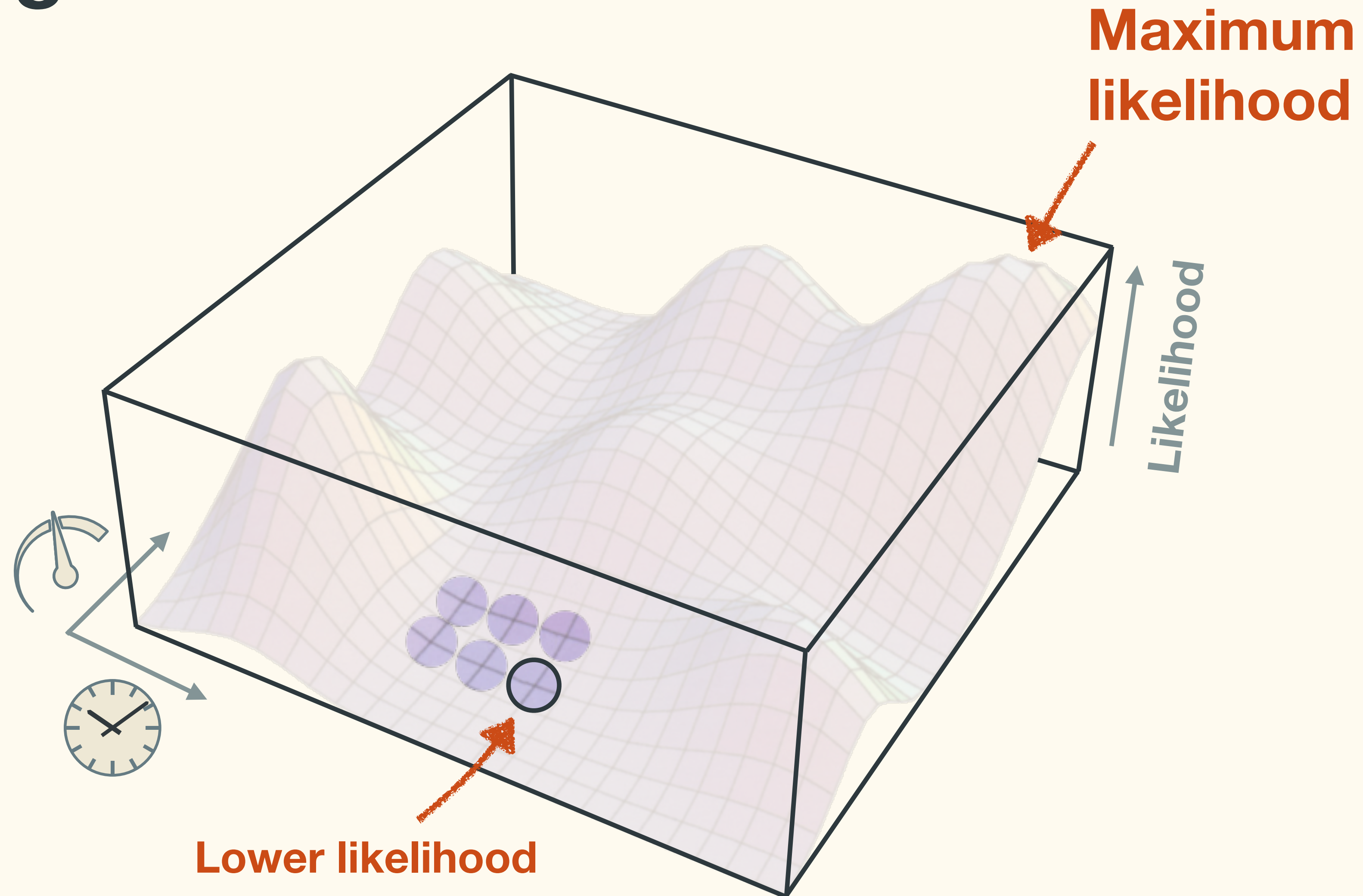
Heuristic search

Hill climbing



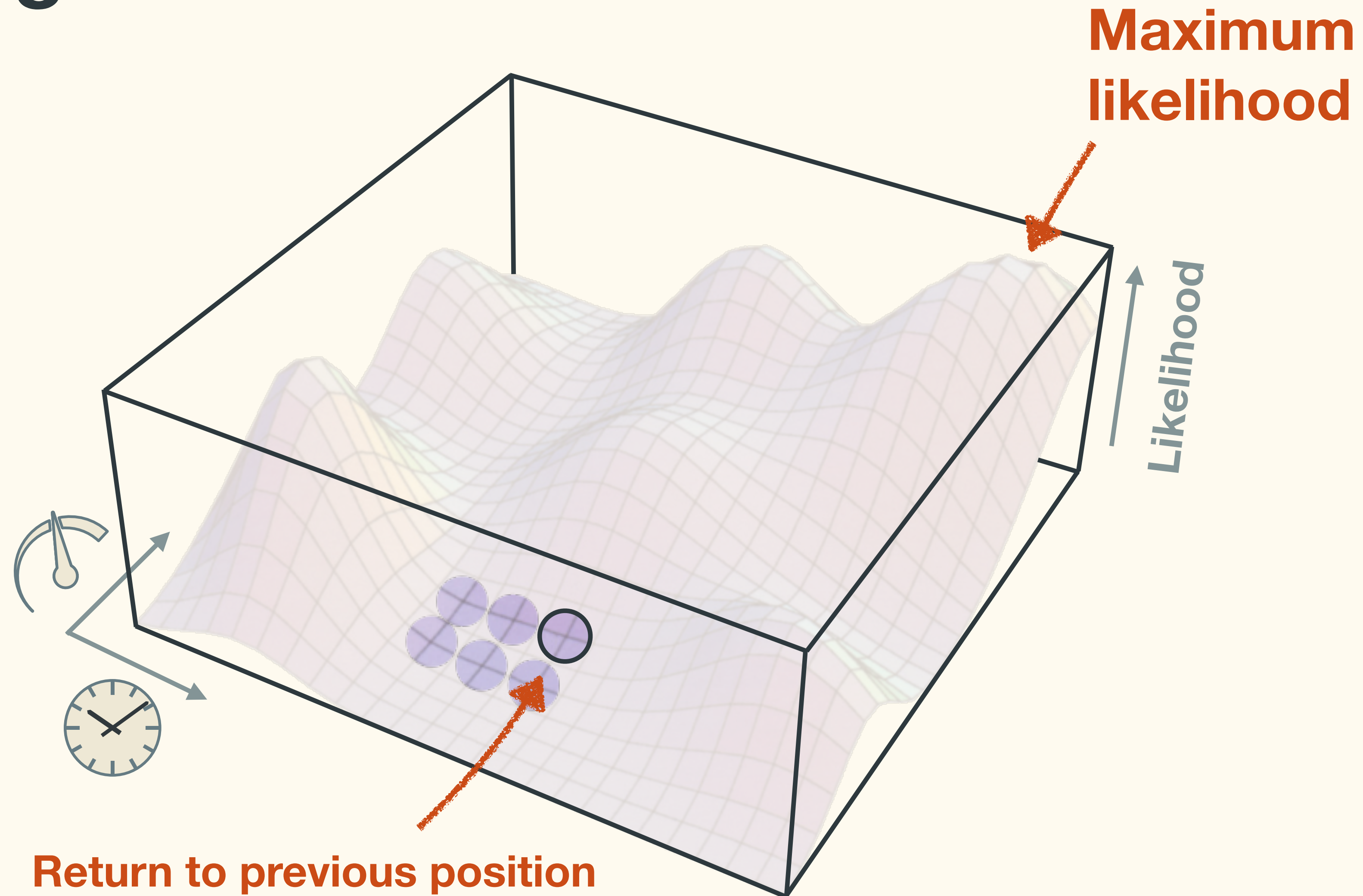
Heuristic search

Hill climbing



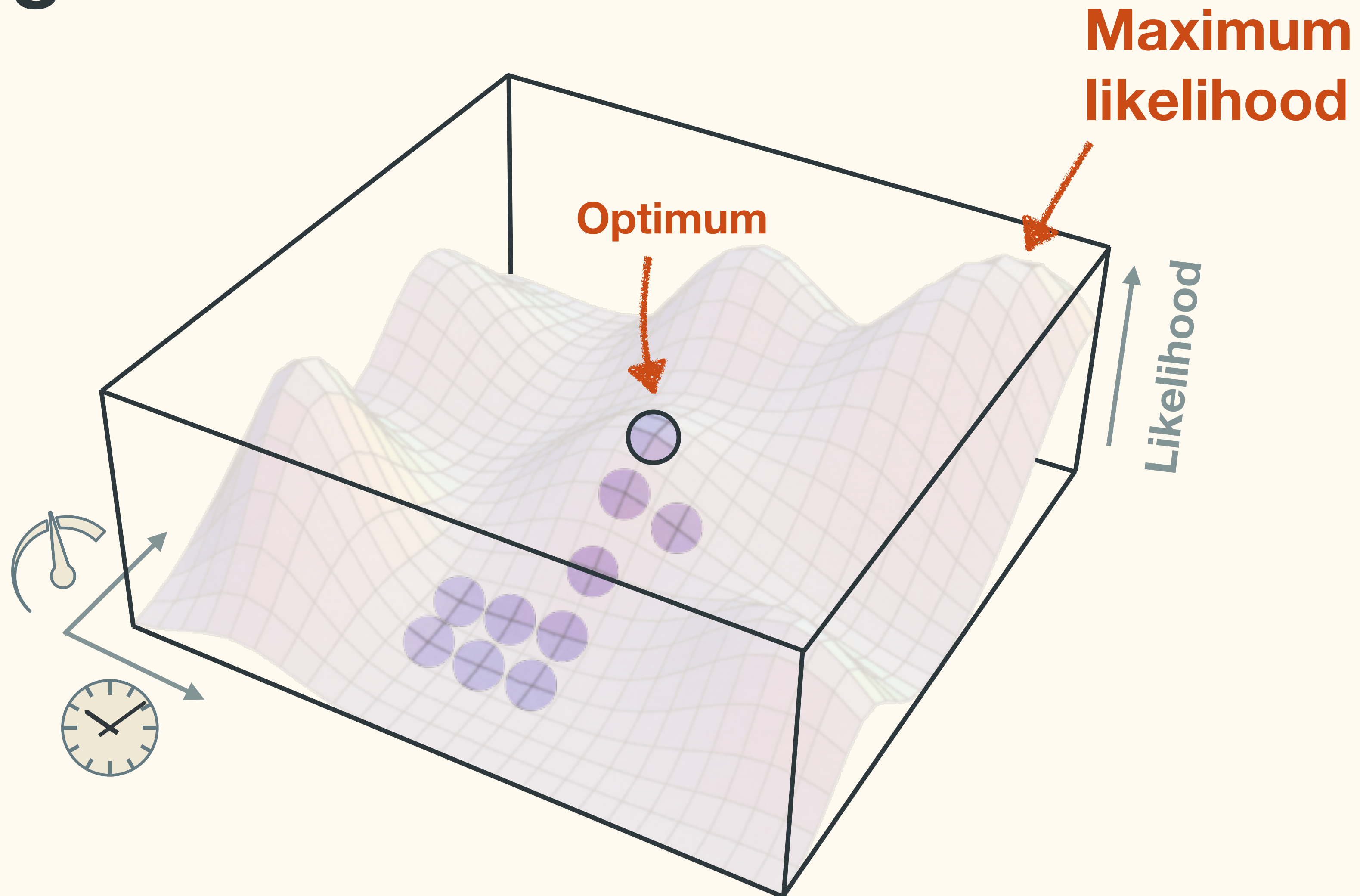
Heuristic search

Hill climbing



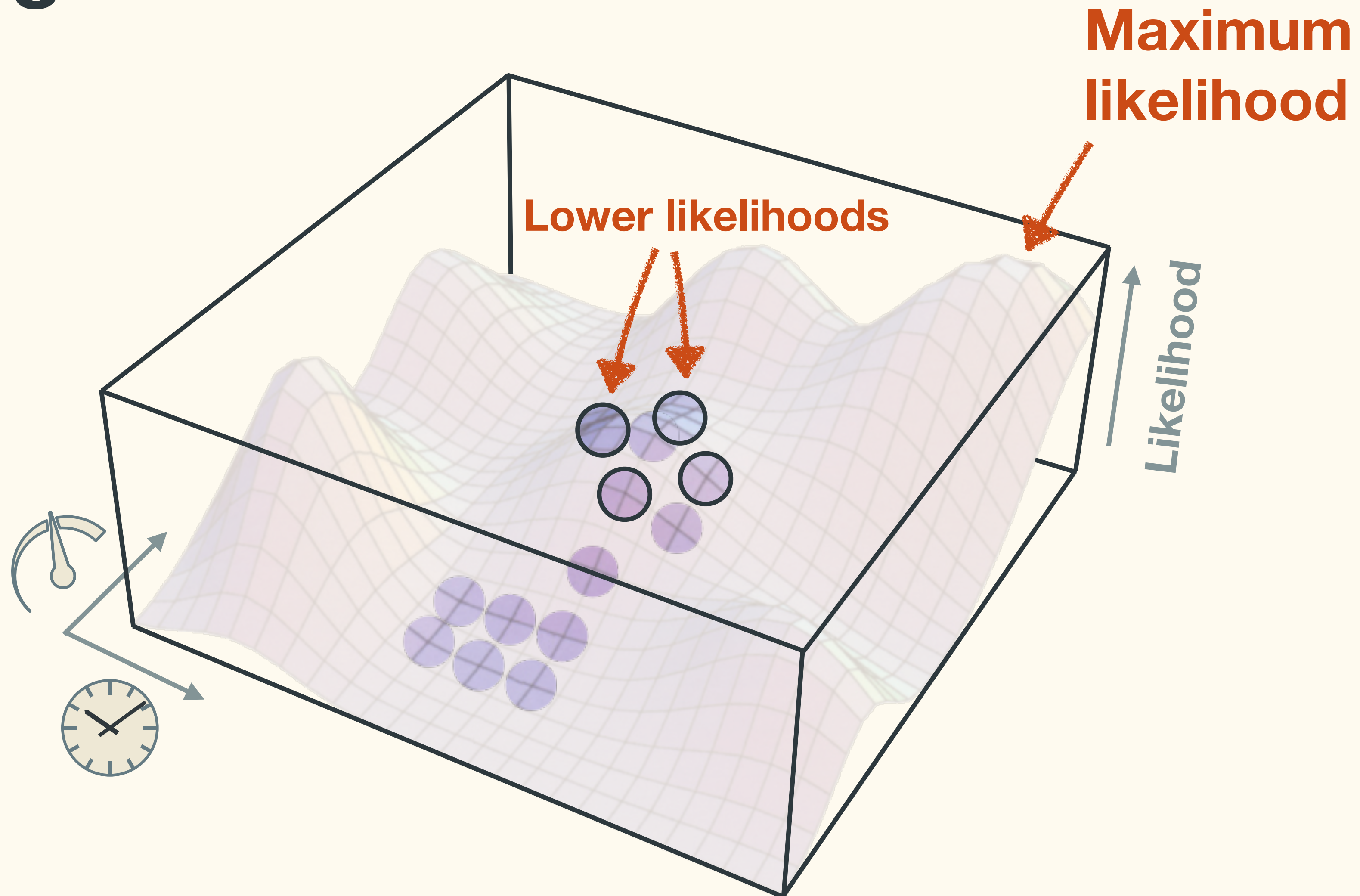
Heuristic search

Hill climbing



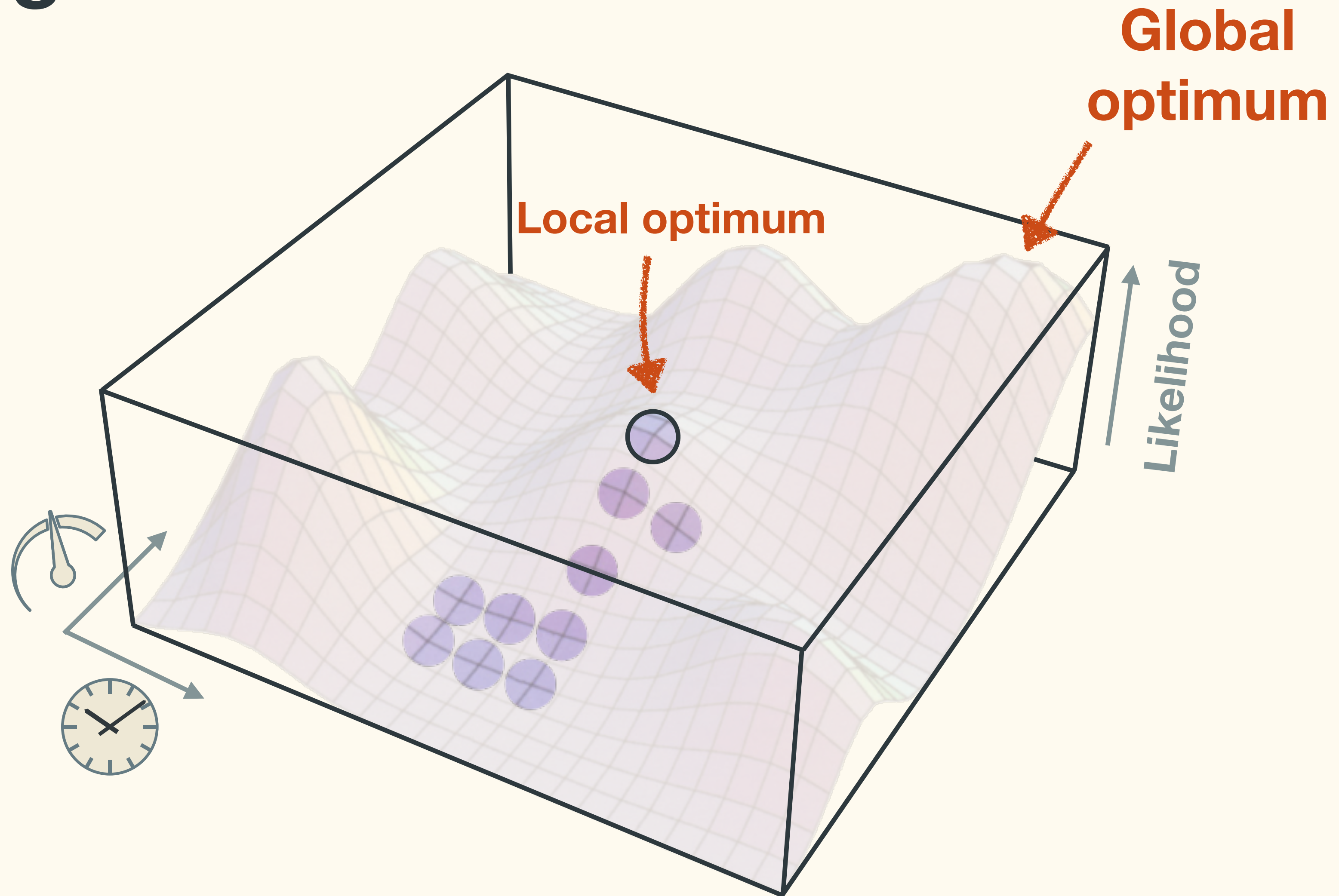
Heuristic search

Hill climbing



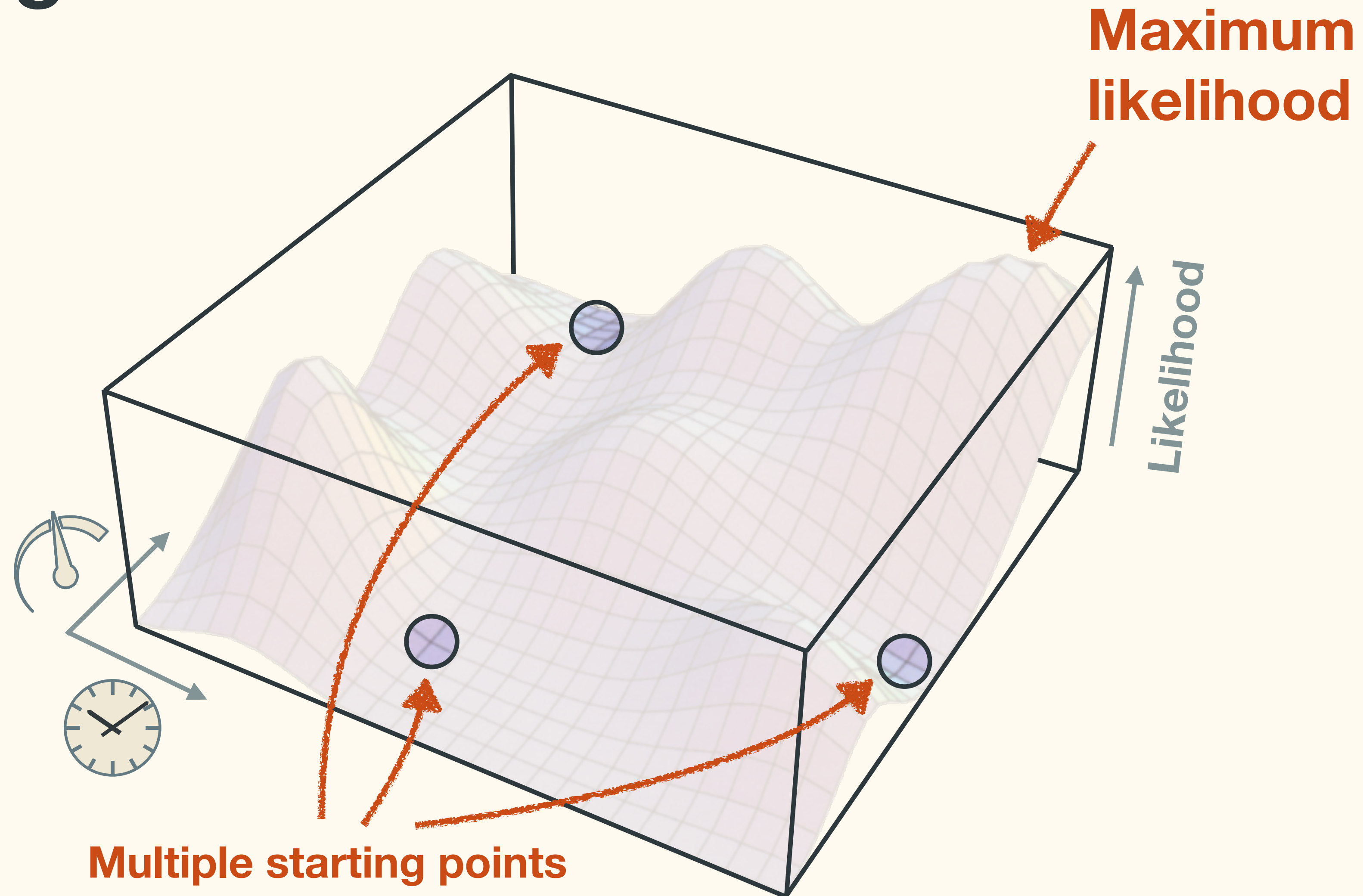
Heuristic search

Hill climbing



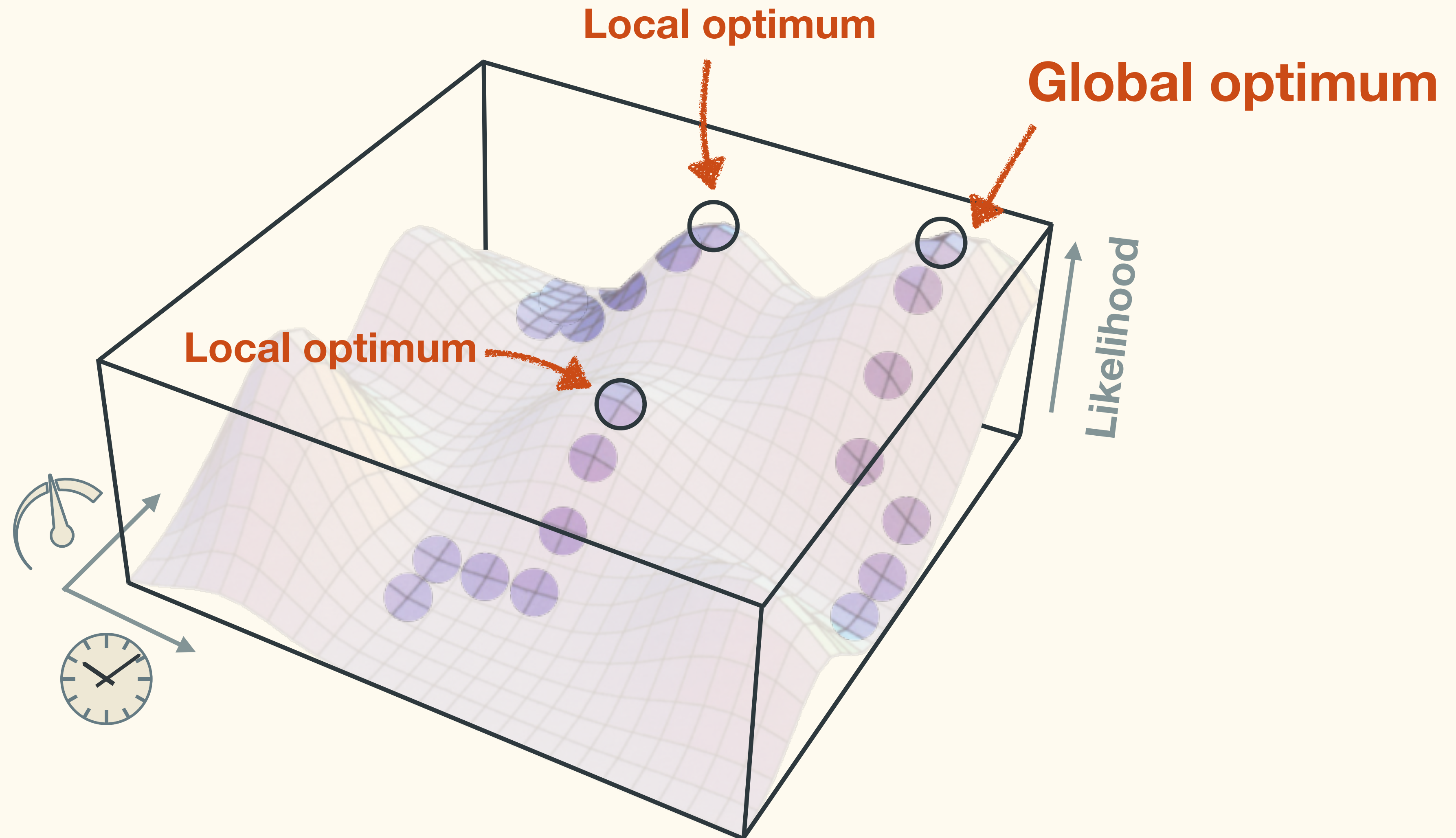
Heuristic search

Hill climbing



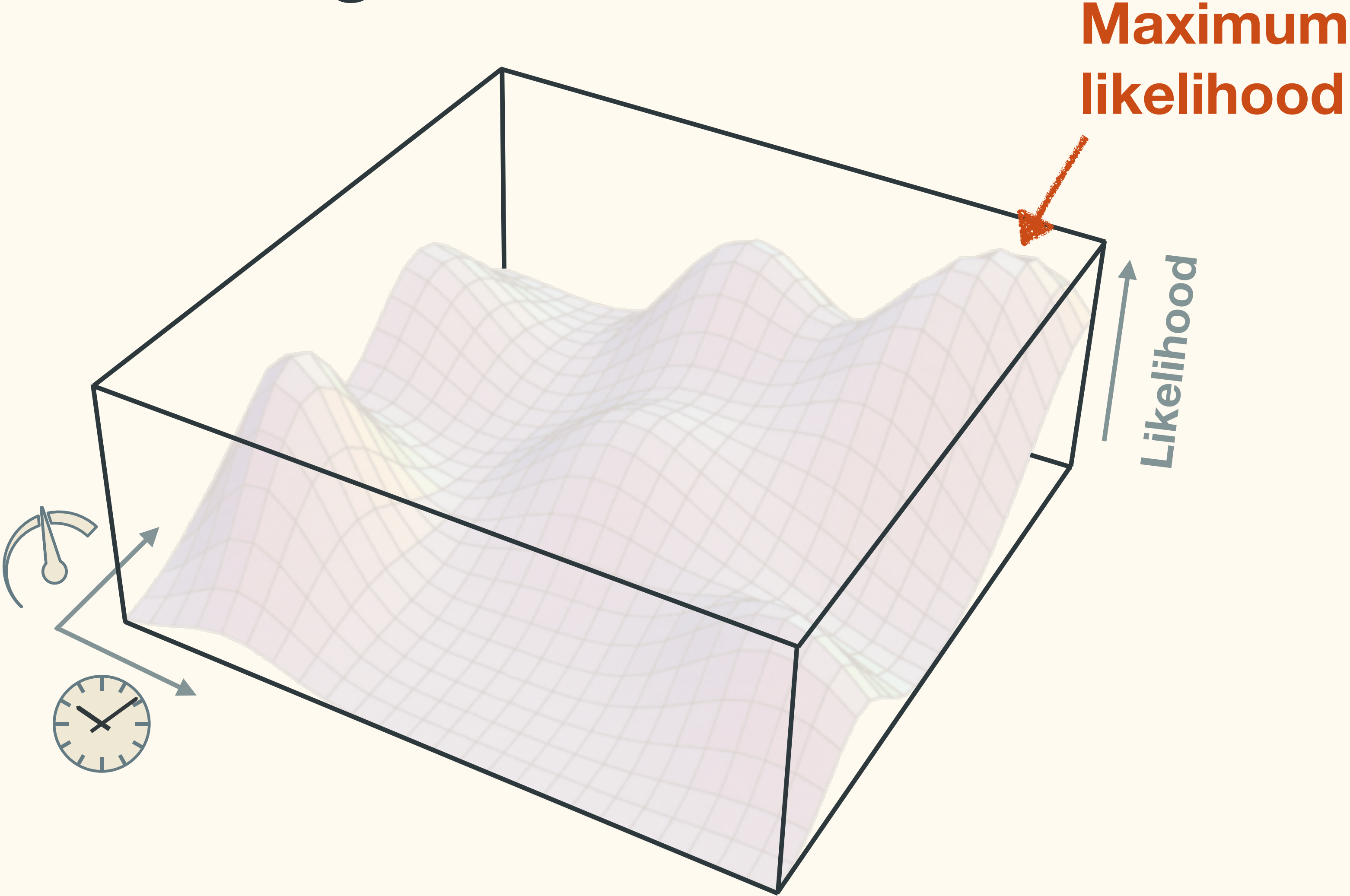
Heuristic search

Hill climbing



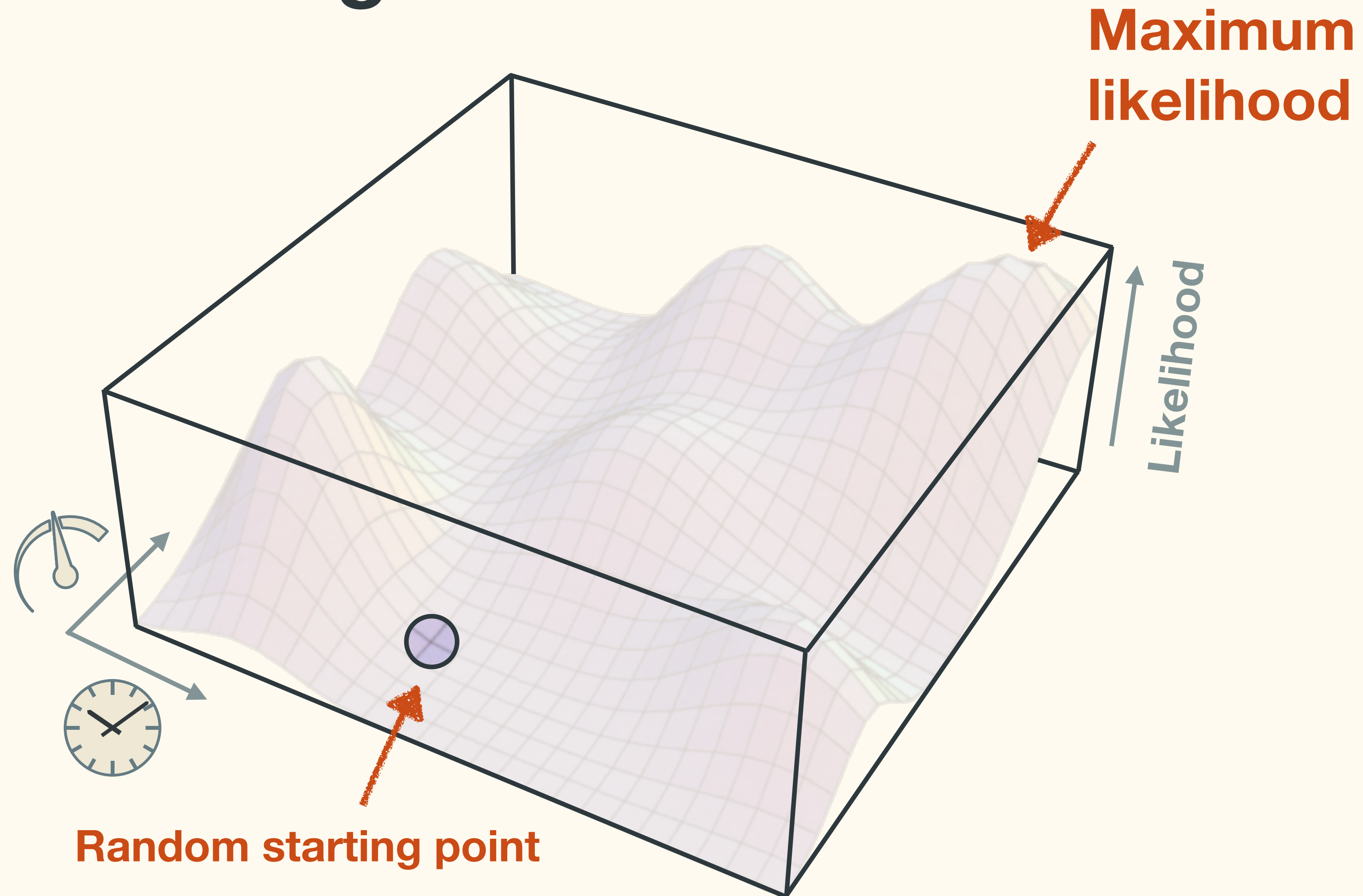
Heuristic search

Simulated annealing



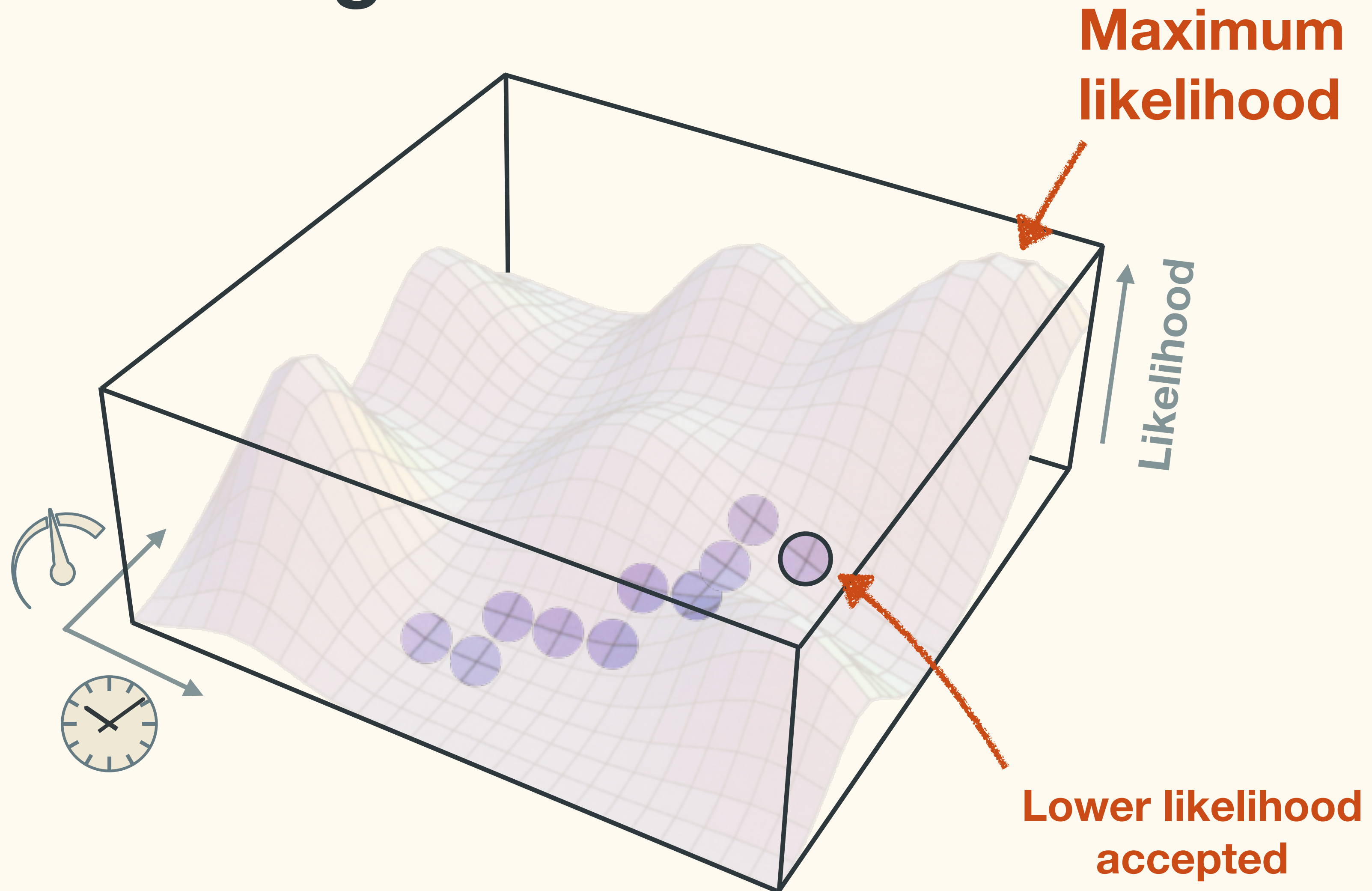
Heuristic search

Simulated annealing



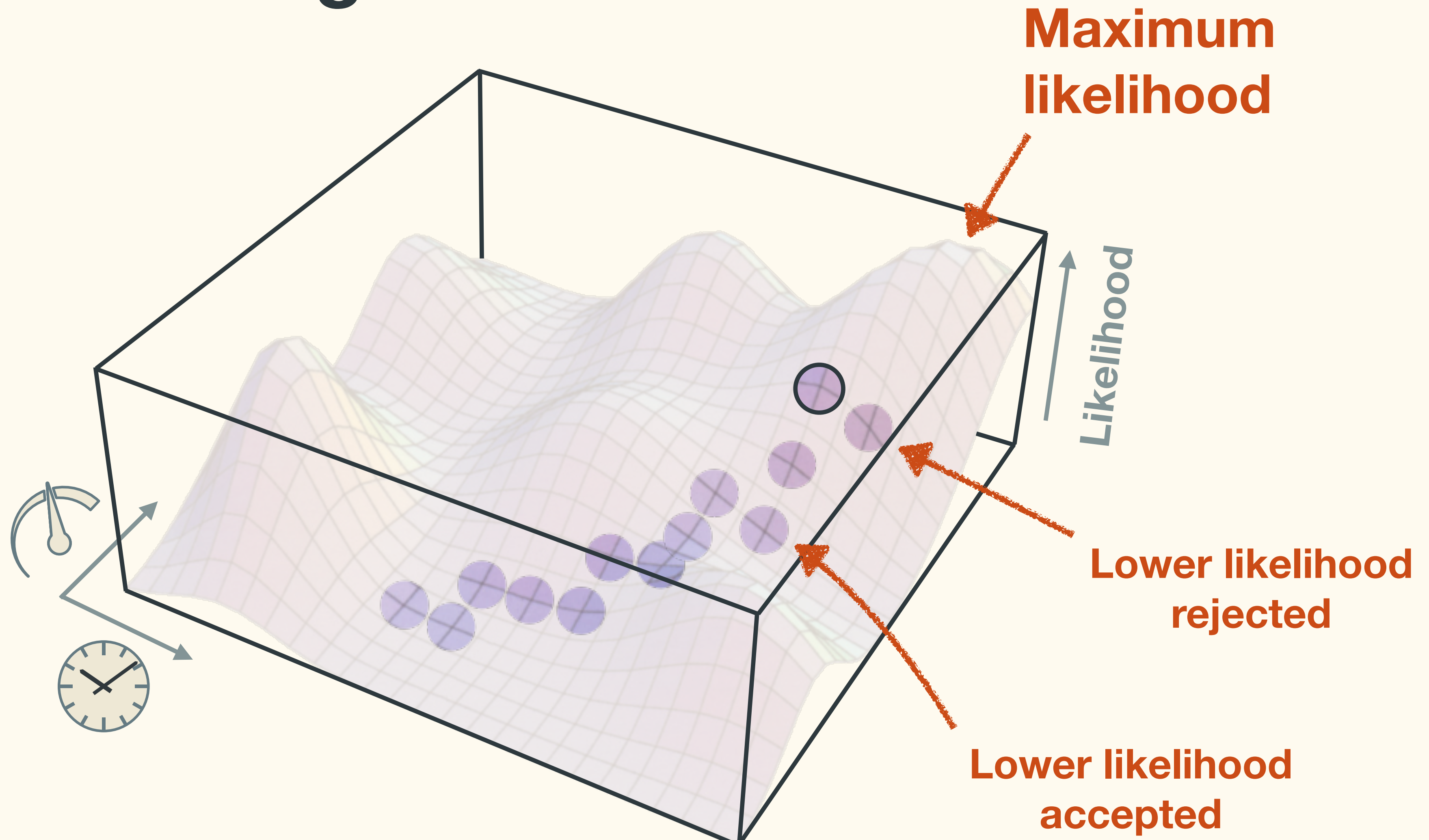
Heuristic search

Simulated annealing



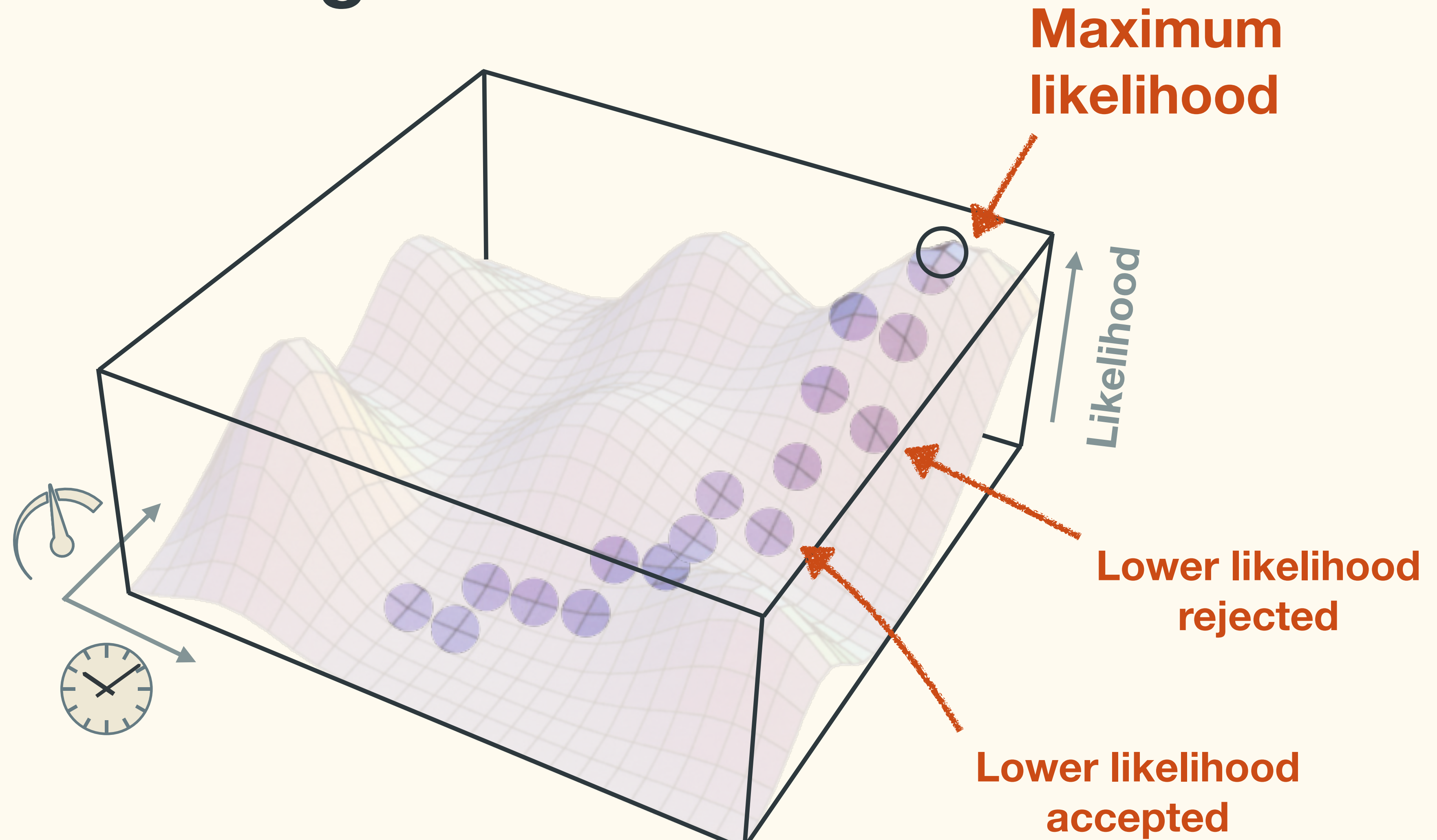
Heuristic search

Simulated annealing



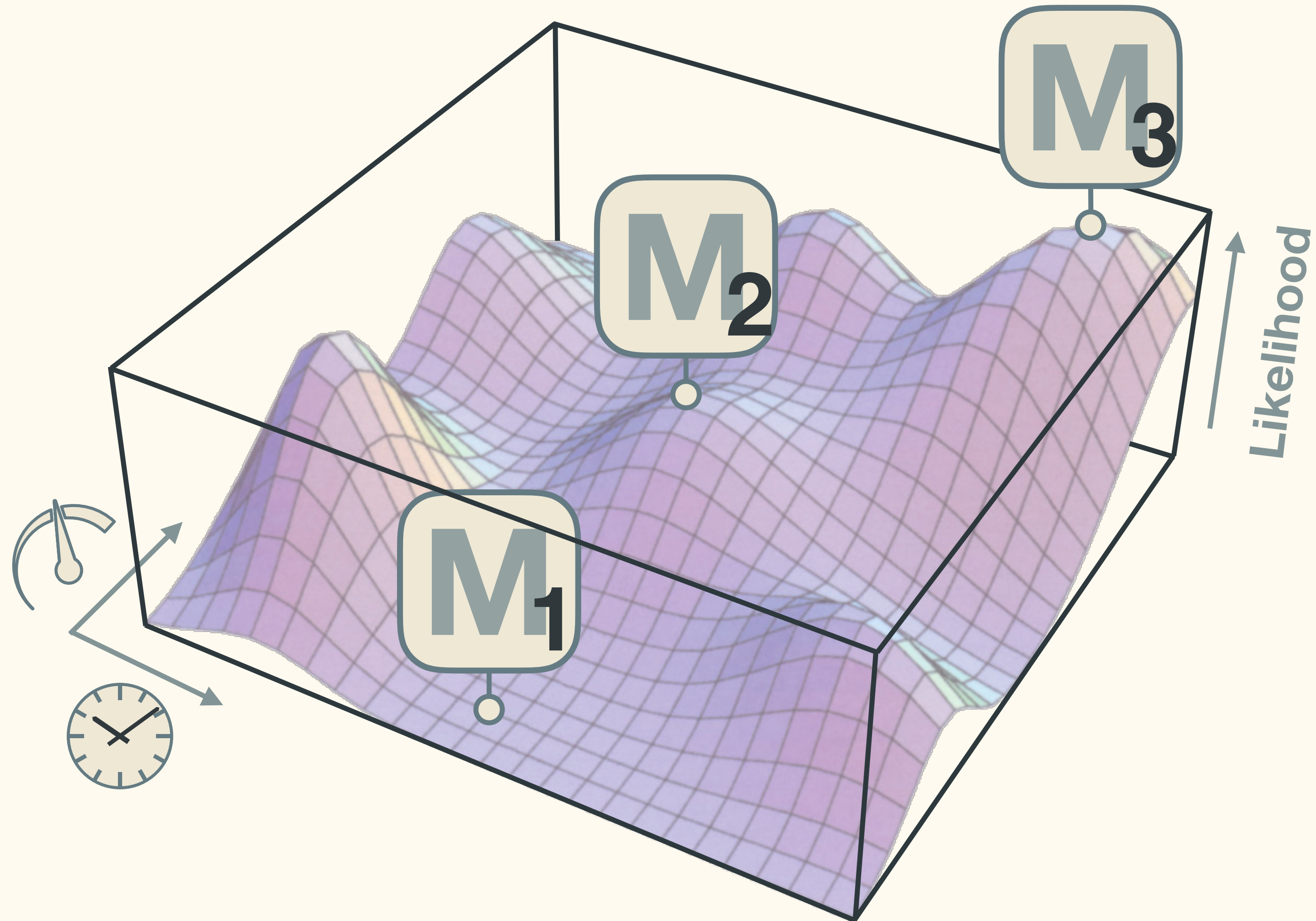
Heuristic search

Simulated annealing



Heuristic search

Simulated annealing



Likelihood

$$\log \left(L \left(\boxed{M_1} \mid \begin{array}{l} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$

$$\log \left(L \left(\boxed{M_2} \mid \begin{array}{l} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -10.3$$

Likelihood



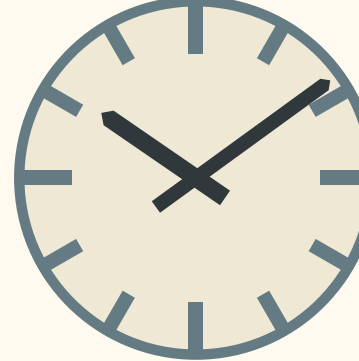
$$\log \left(L \left(\boxed{M_1} \mid \begin{array}{l} \text{ACTTTG} \\ \text{ACTGGG} \end{array} \right) \right) = -13.4$$

$$\log \left(L \left(\boxed{M_2} \mid \begin{array}{l} \text{ACTTTG} \\ \text{ACTGGG} \end{array} \right) \right) = -10.3$$

$$\log \left(L \left(\boxed{M_3} \mid \begin{array}{l} \text{ACTTTG} \\ \text{ACTGGG} \end{array} \right) \right) = -5.4$$

Likelihood

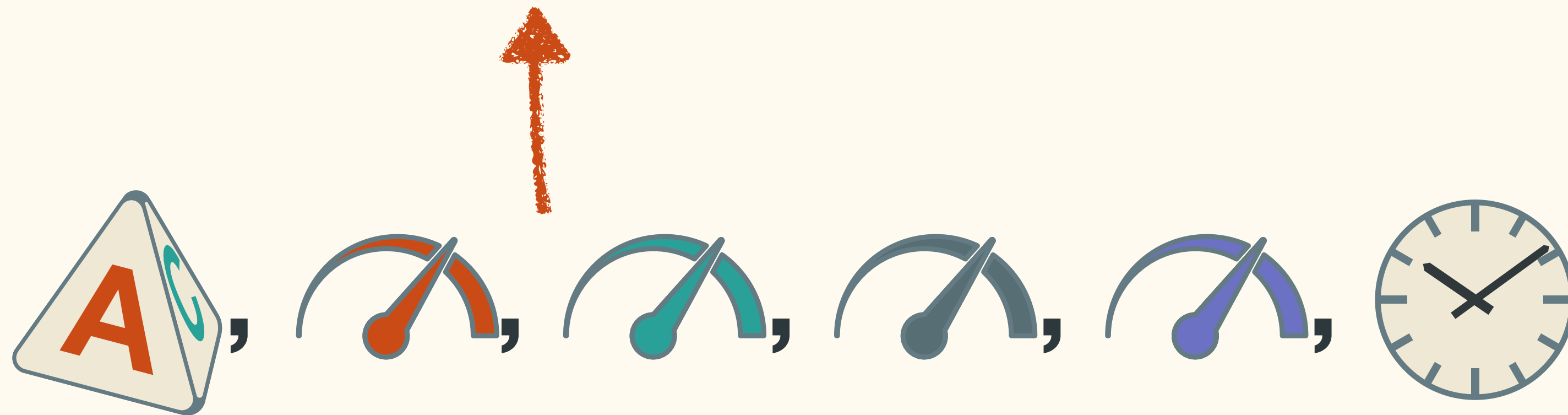
$$\log \left(L \left(\text{M}_3 \mid \begin{array}{c} \text{ACTTGG} \\ \text{ACTGGG} \end{array} \right) \right) = -5.4$$

 ,  = 0.2,  = 0.01

An orange arrow points from the tweezers icon up to the M_3 term in the equation above.

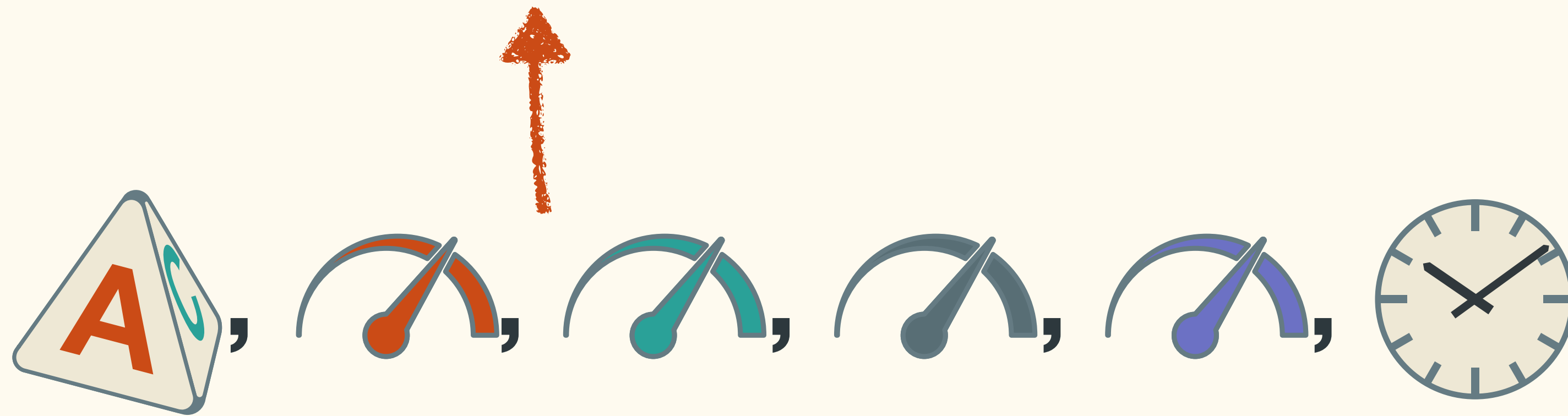
Likelihood

$$\log \left(L \left(\boxed{M}_3 \mid \begin{array}{c} \text{ACTTG} \\ \text{ACTGG} \end{array} \right) \right) = -5.4$$

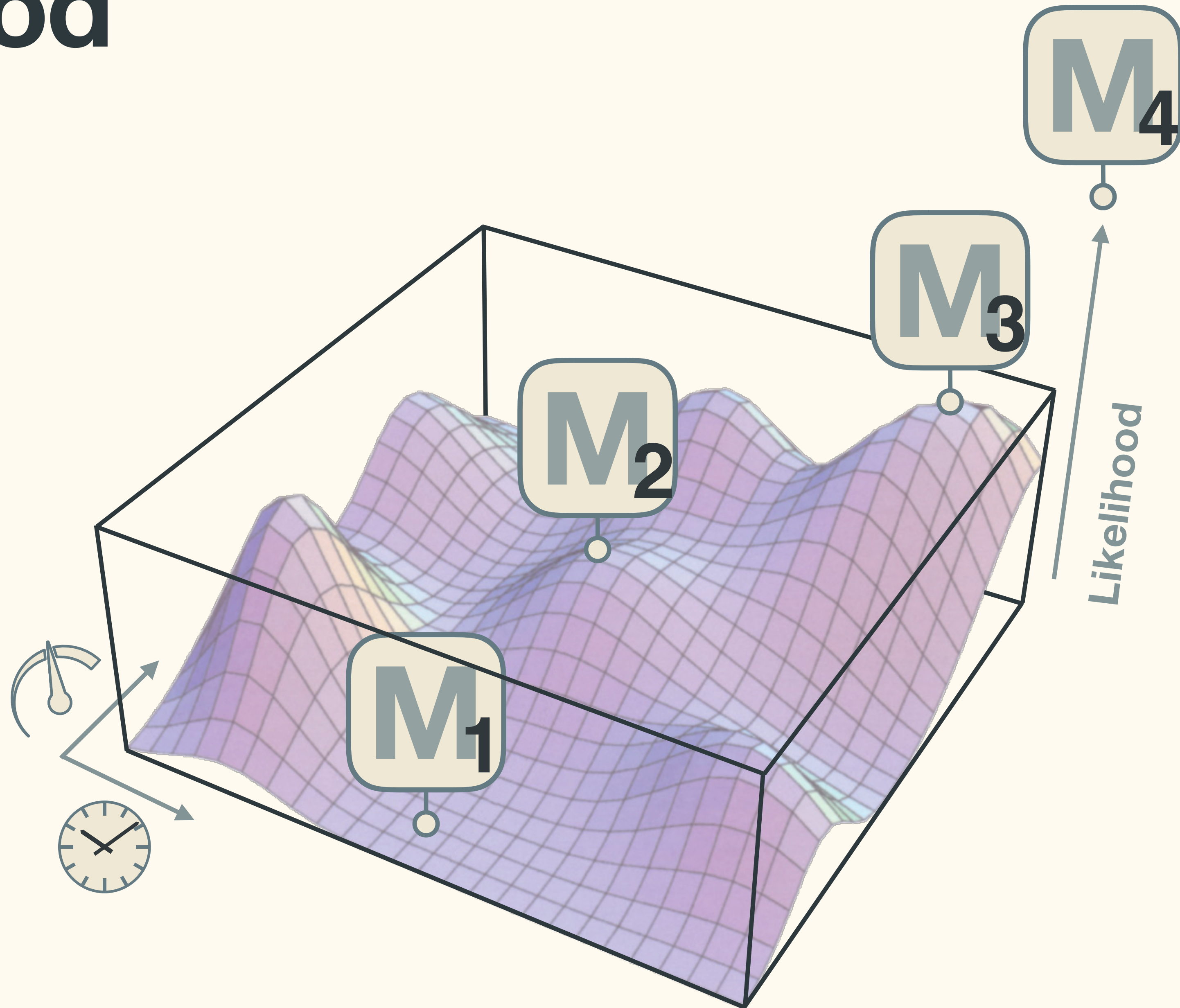


Likelihood

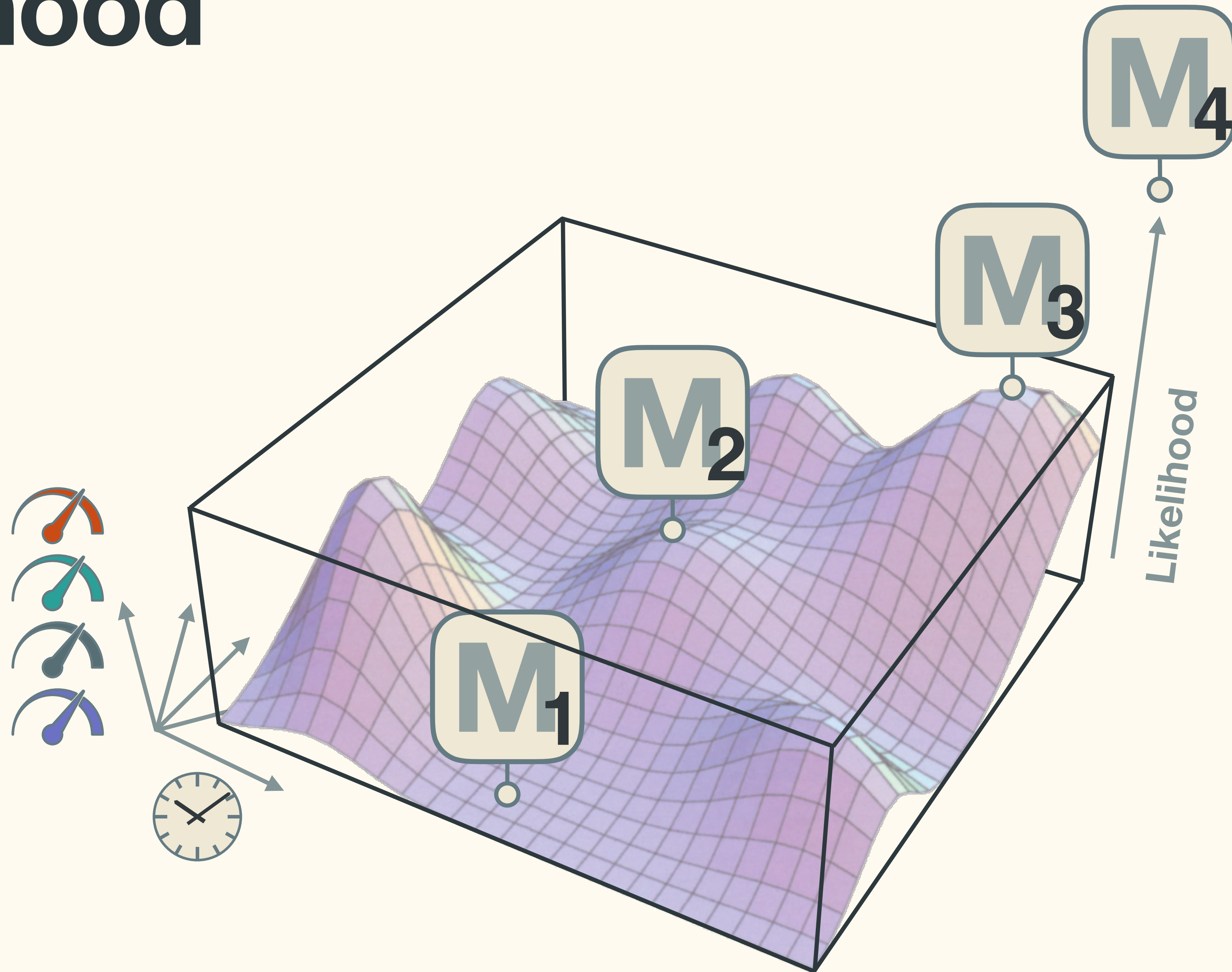
$$\log \left(L \left(\boxed{M}_3 \mid \begin{array}{c} \text{ACTTG} \\ \text{ACTGG} \end{array} \right) \right) = -5.4$$



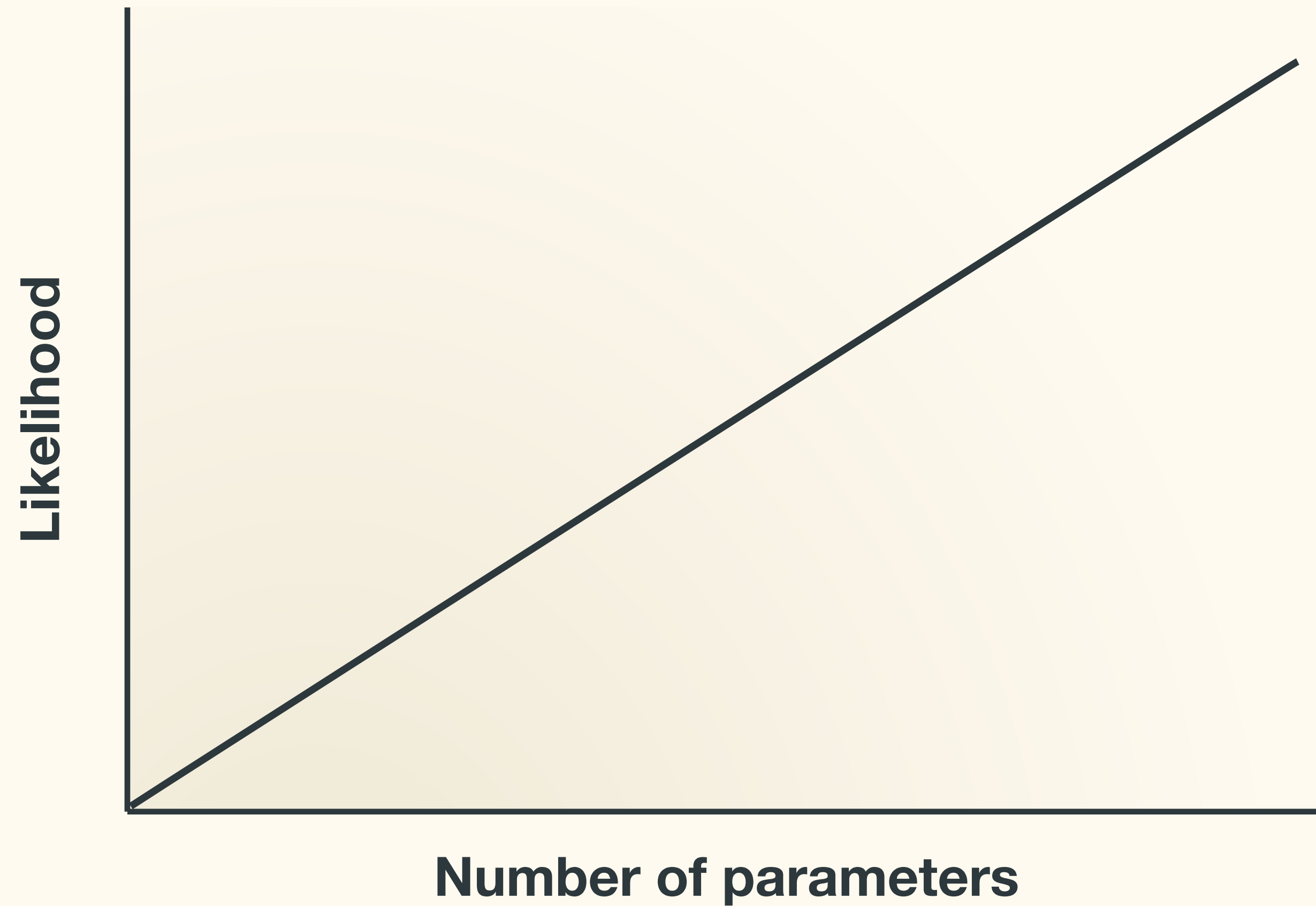
Likelihood



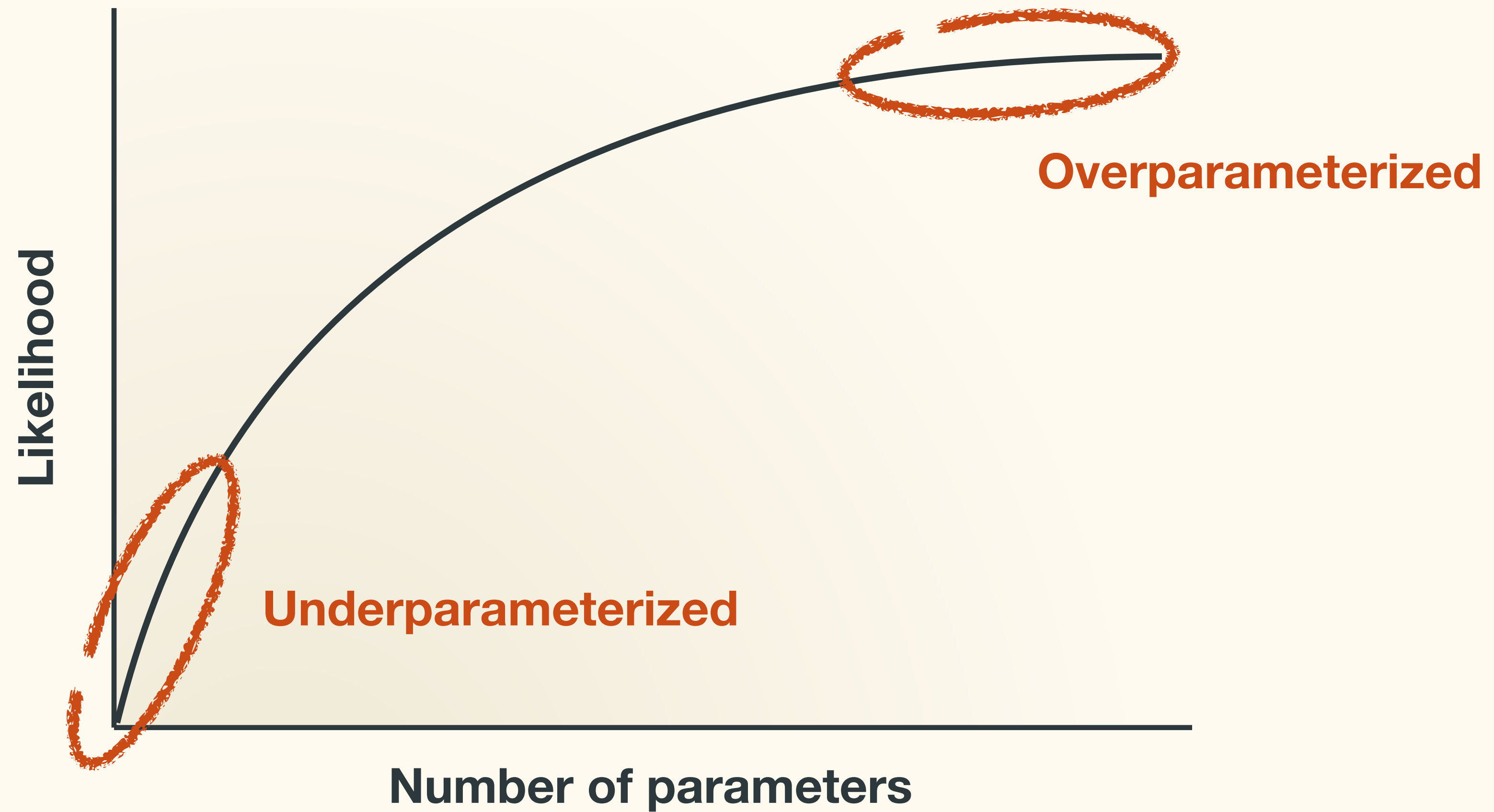
Likelihood



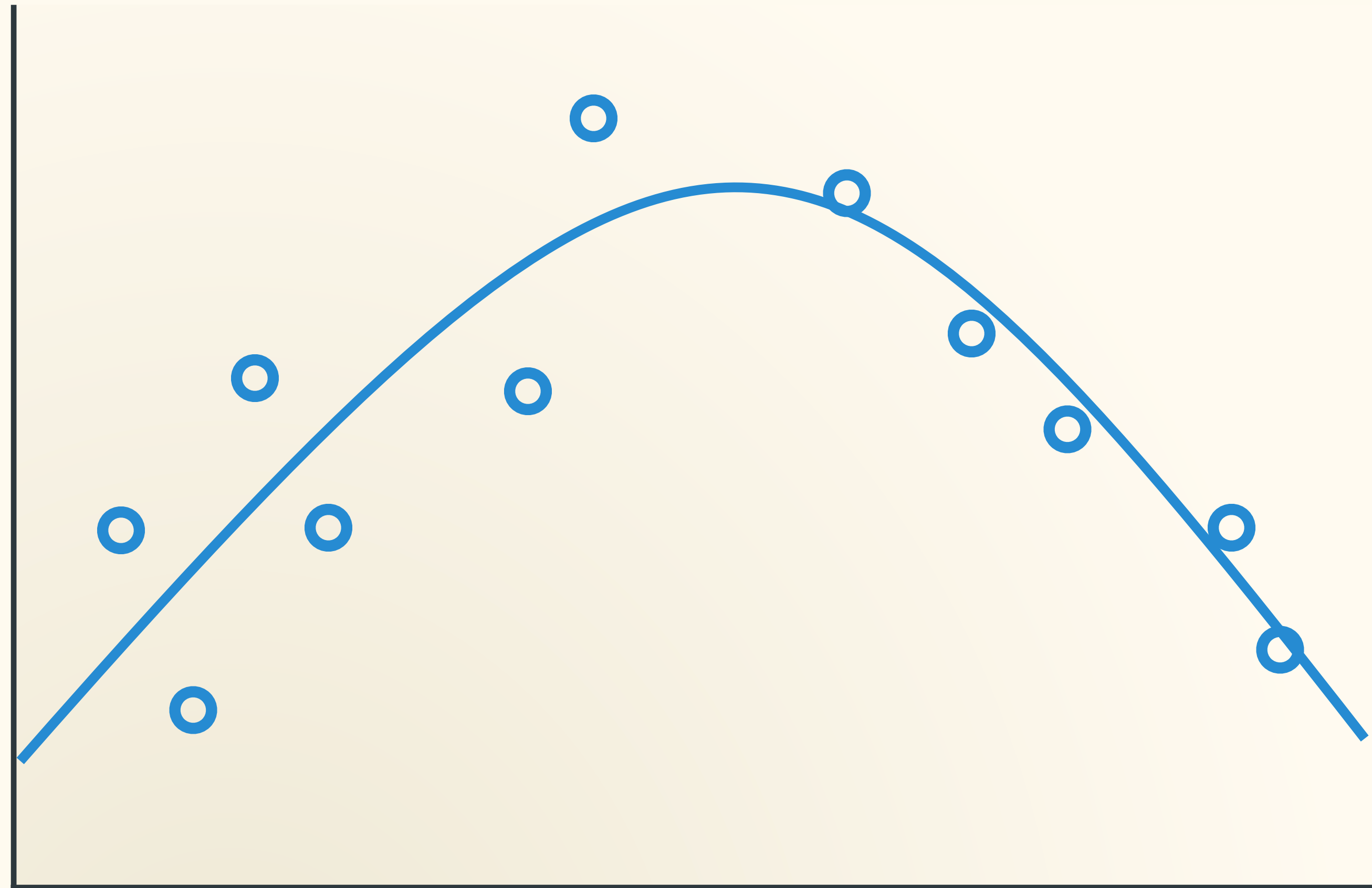
Likelihood



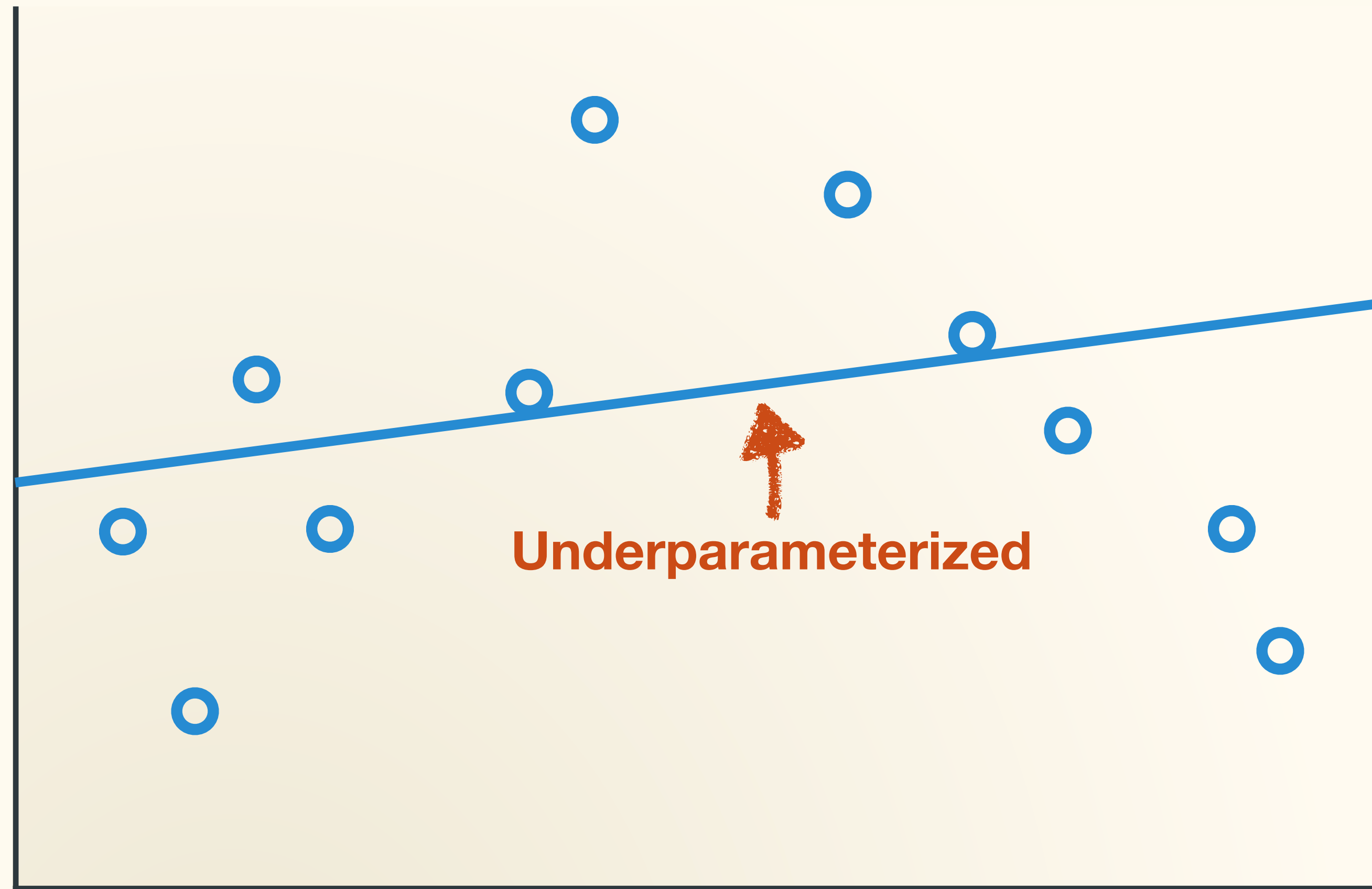
Likelihood



Parameterization



Parameterization



Parameterization



Likelihood ratio test

$$LRT = 2 \log \left(\frac{L \text{ (Complex model)}}{L \text{ (Simple model)}} \right)$$

Likelihood ratio test

$$LRT = 2 \log \left(\frac{M_4}{M_3} \right)$$



Compared to
Chi-square score

Akaike information criterion

$$\text{AIC} = 2k - 2 \left(\log(L) \right)$$

Number of parameters

Akaike information criterion

$$\text{AIC} (M_4) = 2k - 2 \left(\log (L | M_4) \right)$$

Akaike information criterion

$$\text{AIC} (M_4) = 2k - 2 \left(\log (L | M_4) \right)$$

$$\text{AIC} (M_3) = 2k - 2 \left(\log (L | M_3) \right)$$

Akaike information criterion

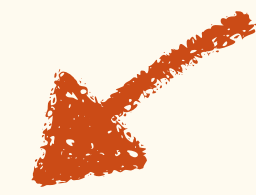
$$\delta AIC = AIC (M_4) - AIC (M_3)$$

Bayesian inference

Bayesian inference

- **A T-rex outside the door?**
- **Somebody pretending to be a T-rex?**

Likelihood



- **A T-rex outside the door?**
- **Somebody pretending to be a T-rex?**



Likelihood

$$L(\text{[T-Rex]} \mid \text{ROAR}) = P(\text{ROAR} \mid \text{[T-Rex]})$$

$$L(\text{[T-Rex Blue]} \mid \text{ROAR}) = P(\text{ROAR} \mid \text{[T-Rex Blue]})$$

Likelihood

$$P(\text{ROAR} \mid \text{Image of a brown T-Rex}) \approx 1$$

$$P(\text{ROAR} \mid \text{Image of a blue T-Rex}) \approx 1$$

Likelihood

$$L \left(\text{[Image of a brown T-Rex]} \mid \text{ROAR} \right) \approx 1$$

$$L \left(\text{[Image of a blue T-Rex]} \mid \text{ROAR} \right) \approx 1$$

Probability

$$P \left(\text{[Tyrannosaurus Rex]} \mid \text{ROAR} \right) = ?$$

Bayesian inference



Thomas Bayes
(1701–1761)

LII. *An Essay towards solving a Problem in the Doctrine of Chances.* By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.


Dear Sir,

Read Dec. 23, 1763. I Now send you an essay which I have found among the papers of our deceased friend Mr. Bayes, and which, in my opinion, has great merit, and well deserves to be preserved. Experimental philosophy, you will find, is nearly interested in the subject of it; and on this account there seems to be particular reason for thinking that a communication of it to the Royal Society cannot be improper.

He had, you know, the honour of being a member of that illustrious Society, and was much esteemed by many in it as a very able mathematician. In an introduction which he has writ to this Essay, he says, that his design at first in thinking on the subject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circum-

Bayes' theorem

Model


$$P(\text{Dinosaur} \mid \text{ROAR}) = \frac{P(\text{ROAR} \mid \text{Dinosaur}) \times P(\text{Dinosaur})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{🦖} \mid \text{ROAR}) = \frac{P(\text{ROAR} \mid \text{🦖}) \times P(\text{🦖})}{P(\text{ROAR})}$$

The diagram illustrates Bayes' theorem for a classification task. The left side of the equation represents the posterior probability, $P(\text{🦖} \mid \text{ROAR})$, where the dinosaur icon is enclosed in a rounded square. An orange arrow labeled "Data" points from the word "ROAR" to the dinosaur icon. The right side of the equation shows the numerator as the product of the likelihood, $P(\text{ROAR} \mid \text{🦖})$, and the prior probability, $P(\text{🦖})$. Both the likelihood and prior terms have the dinosaur icon enclosed in a rounded square. The denominator is the marginal probability, $P(\text{ROAR})$.

Bayes' theorem

Posterior probability

$$P(\text{ROAR} \mid \text{Dinosaur}) = \frac{P(\text{Dinosaur} \mid \text{ROAR}) \times P(\text{Dinosaur})}{P(\text{ROAR})}$$

The image shows the Bayesian formula for the probability of a dinosaur given a roar. The posterior probability $P(\text{Dinosaur} \mid \text{ROAR})$ is circled in orange. The numerator consists of the likelihood $P(\text{Dinosaur} \mid \text{ROAR})$ and the prior $P(\text{Dinosaur})$. The denominator is the marginal probability $P(\text{ROAR})$. The word "ROAR" is written in orange, and the dinosaur image is in a blue box.

Bayes' theorem

$$P(\text{🦖} \mid \text{ROAR}) = \frac{\overset{\text{Likelihood}}{L(\text{🦖} \mid \text{ROAR})} \times P(\text{🦖})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{🦖} | \text{ROAR}) = \frac{L(\text{🦖} | \text{ROAR}) \times P(\text{🦖})}{P(\text{ROAR})}$$

Prior probability

Bayes' theorem

$$P(\text{🦖} | \text{ROAR}) = \frac{L(\text{🦖} | \text{ROAR}) \times P(\text{🦖})}{P(\text{ROAR})}$$

Constant

Bayes' theorem

Posterior probability

Likelihood

Prior probability

$$P\left(\text{[Tyrannosaurus Rex]} \mid \text{ROAR}\right) = \frac{L\left(\text{[Tyrannosaurus Rex]} \mid \text{ROAR}\right) \times P\left(\text{[Tyrannosaurus Rex]}\right)}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{\overset{\approx 1}{L(\text{Brown T-Rex} \mid \text{ROAR})} \times P(\text{Brown T-Rex})}{P(\text{ROAR})}$$

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{L(\text{Blue T-Rex} \mid \text{ROAR}) \times P(\text{Blue T-Rex})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{1}{P(\text{ROAR})} \times P(\text{Brown T-Rex})$$

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{\overset{\approx 1}{L(\text{Blue T-Rex} \mid \text{ROAR})}}{P(\text{ROAR})} \times P(\text{Blue T-Rex})$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{1}{P(\text{ROAR})} \times P(\text{Brown T-Rex})$$

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{1}{P(\text{ROAR})} \times P(\text{Blue T-Rex})$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{P(\text{Brown T-Rex})}{P(\text{ROAR})}$$

Prior probabilities

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{P(\text{Blue T-Rex})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{P(\text{Brown T-Rex})}{P(\text{ROAR})} \approx 0.00000001$$

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{P(\text{Blue T-Rex})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{🦖} \mid \text{ROAR}) = \frac{P(\text{🦖})}{P(\text{ROAR})} \approx \frac{0.00000001}{P(\text{ROAR})}$$

$$P(\text{🦖} \mid \text{ROAR}) = \frac{P(\text{🦖})}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{Brown T-Rex} \mid \text{ROAR}) = \frac{P(\text{Brown T-Rex})}{P(\text{ROAR})} \approx \frac{0.00000001}{P(\text{ROAR})}$$

$$P(\text{Blue T-Rex} \mid \text{ROAR}) = \frac{P(\text{Blue T-Rex})}{P(\text{ROAR})} \approx 0.1$$

Bayes' theorem

$$P(\text{🦖} \mid \text{ROAR}) = \frac{P(\text{🦖})}{P(\text{ROAR})} \approx \frac{0.00000001}{P(\text{ROAR})}$$

$$P(\text{🦖🦕} \mid \text{ROAR}) = \frac{P(\text{🦖🦕})}{P(\text{ROAR})} \approx \frac{0.1}{P(\text{ROAR})}$$

Bayes' theorem

$$P(\text{Blue T-Rex} \mid \text{ROAR}) \gg P(\text{Brown T-Rex} \mid \text{ROAR})$$

Likelihood

Fair dice

$$L(\text{Fair dice} \mid \text{3, 3, 3}) = 1/6$$

Trick dice

$$L(\text{Trick dice} \mid \text{3, 3, 3}) = 1$$

Likelihood

Fair dice

$$L(\text{Fair dice} \mid \text{3 dots}) = 1/6$$

Trick dice

$$L(\text{Trick dice} \mid \text{3 dots}) = 1$$

Bayes' theorem

Posterior probability Likelihood Prior probability

$$P(\text{die} \mid \text{roll}) = \frac{L(\text{die} \mid \text{roll}) \times P(\text{die})}{P(\text{roll})}$$
$$P(\text{die} \mid \text{roll}) = \frac{L(\text{die} \mid \text{roll}) \times P(\text{die})}{P(\text{roll})}$$

The diagram illustrates Bayes' theorem for a die roll. It shows two identical equations. The first equation is labeled with 'Posterior probability' above the left side, 'Likelihood' above the numerator's first term, and 'Prior probability' above the numerator's second term. The second equation is identical but lacks these labels. In both equations, the left side represents the posterior probability of a die given a roll of 3. The numerator consists of the likelihood of a die given a roll of 3, multiplied by the prior probability of the die. The denominator is the marginal probability of a roll of 3. The die and roll are represented by icons: a 3D die and a 2D die face showing 3 dots.

Bayes' theorem

$$P(\text{die} \mid \text{roll}) = \frac{L(\text{die} \mid \text{roll}) \times P(\text{die})}{P(\text{roll})} = \frac{1}{6}$$

The equation above shows the calculation of the posterior probability for a 3-sided die. The likelihood term $L(\text{die} \mid \text{roll})$ is circled in orange, and the value $= 1/6$ is written above it.

$$P(\text{die} \mid \text{roll}) = \frac{L(\text{die} \mid \text{roll}) \times P(\text{die})}{P(\text{roll})}$$

The equation below shows the general form of Bayes' theorem for a 6-sided die, where the likelihood term is not circled.

Bayes' theorem

$$P(\text{die} \mid \text{face}) = \frac{1/6 \times P(\text{die})}{P(\text{face})}$$

$$P(\text{die} \mid \text{face}) = \frac{L(\text{die} \mid \text{face}) \times P(\text{die})}{P(\text{face})}$$

Bayes' theorem

$$P(\text{die} \mid \text{roll}) = \frac{1/6 \times P(\text{die})}{P(\text{roll})}$$

$$P(\text{die} \mid \text{roll}) = \frac{L(\text{die} \mid \text{roll}) \times P(\text{die})}{P(\text{roll})} = 1$$

Bayes' theorem

$$P(\text{die} \mid \text{two}) = \frac{1/6 \times P(\text{die})}{P(\text{two})}$$

$$P(\text{cube} \mid \text{two}) = \frac{1 \times P(\text{cube})}{P(\text{two})}$$

Bayes' theorem

$$P(\text{die} \mid \text{two}) = \frac{1/6 \times P(\text{die})}{P(\text{two})}$$

$$P(\text{die} \mid \text{two}) = \frac{P(\text{die})}{P(\text{two})}$$

Bayes' theorem

$$P(\text{die} \mid \text{two}) = \frac{1/6 \times P(\text{die})}{P(\text{two})} = 0.99999$$

$$P(\text{die} \mid \text{two}) = \frac{P(\text{die})}{P(\text{two})}$$

Bayes' theorem

$$P(\text{die} \mid \text{two}) = \frac{1/6 \times P(\text{die})}{P(\text{two})} = 0.99999$$

$$P(\text{die} \mid \text{two}) = \frac{P(\text{die})}{P(\text{two})} = 0.00001$$

Bayes' theorem

$$P(\text{die} \mid \text{two dots}) = \frac{1/6 \times P(\text{die})}{P(\text{two dots})} \approx \frac{1/6}{P(\text{two dots})} = 0.99999$$

$$P(\text{die} \mid \text{two dots}) = \frac{P(\text{die})}{P(\text{two dots})} = 0.00001$$

Bayes' theorem

$$P(\text{die} \mid \text{two dots}) = \frac{1/6 \times P(\text{die})}{P(\text{two dots})} \approx \frac{1/6}{P(\text{two dots})} = 0.99999$$

$$P(\text{die} \mid \text{two dots}) = \frac{P(\text{die})}{P(\text{two dots})} = \frac{0.00001}{P(\text{two dots})}$$

Estimating model parameters using Bayesian inference

MCMC

Markov-chain Monte Carlo

Monte Carlo



Monte Carlo methods



Stanisław Ulam
(1909–1984)



Monte Carlo methods



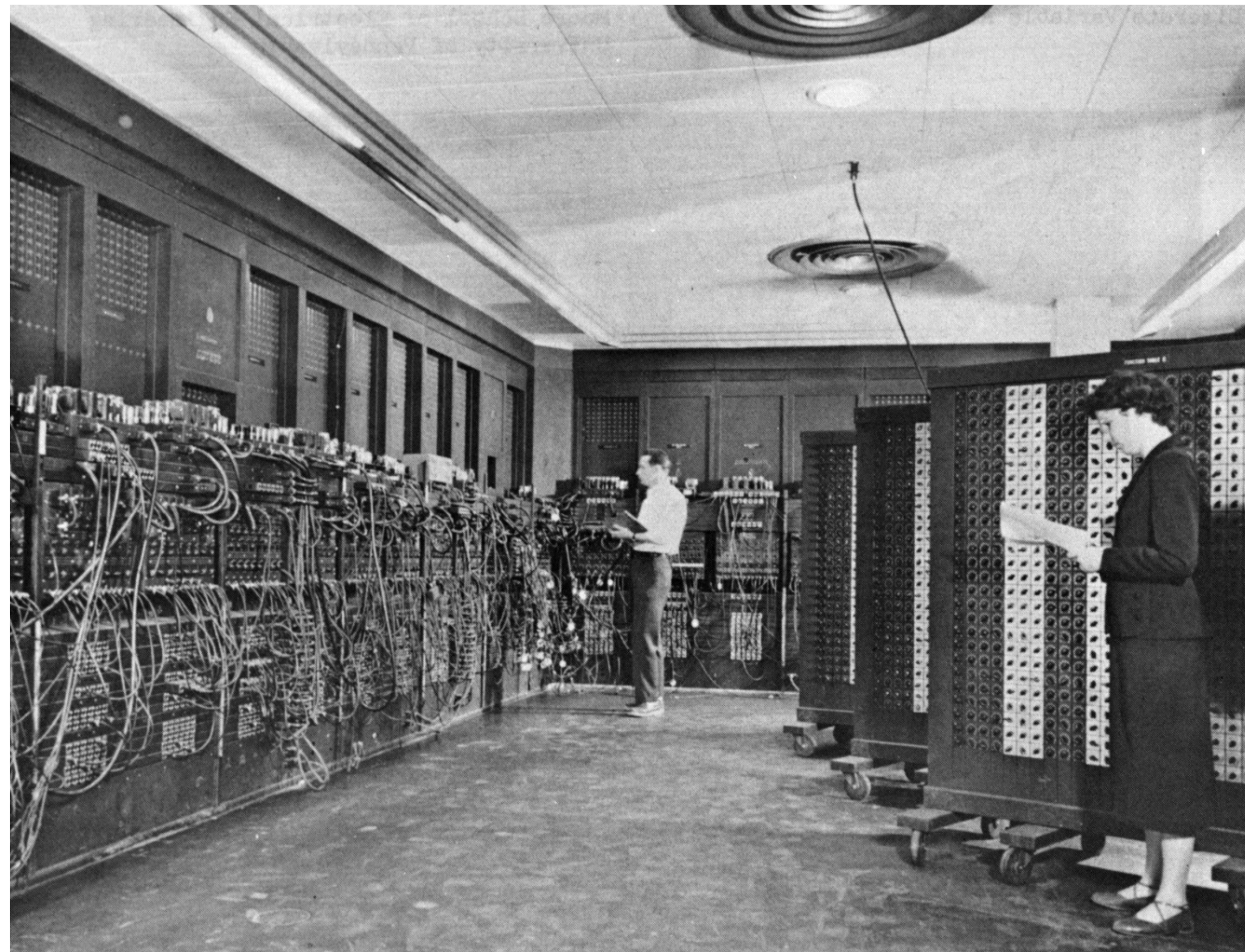
Stanisław Ulam
(1909–1984)

*“What are the chances
that a Canfield solitaire
laid out with 52 cards will
come out successfully?”*

Stanisław Ulam, 1946

Monte Carlo methods

ENIAC
1946



Monte Carlo methods



Stanisław Ulam
(1909–1984)

“Stan had an uncle who would borrow money from relatives because he ‘just had to go to Monte Carlo’.”

Nicholas Metropolis

Monte Carlo



Markov chains



Andrey Markov
(1856–1922)

Markov chains



Markov chains



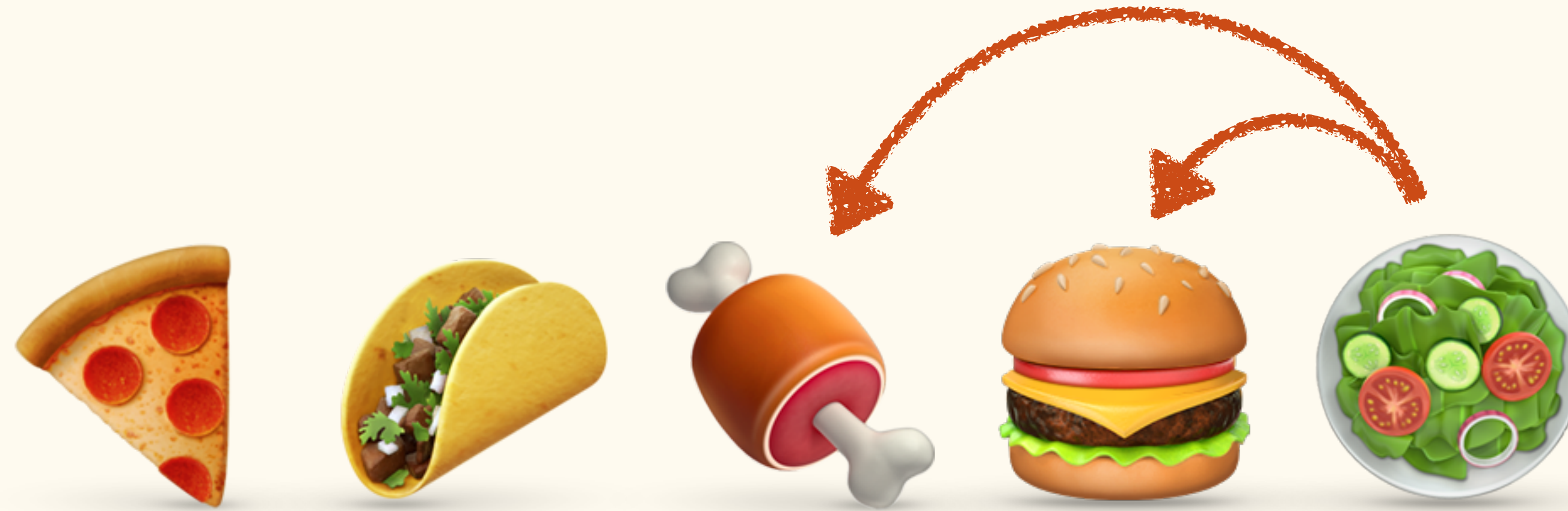
Markov chains



Markov chains



Markov chains



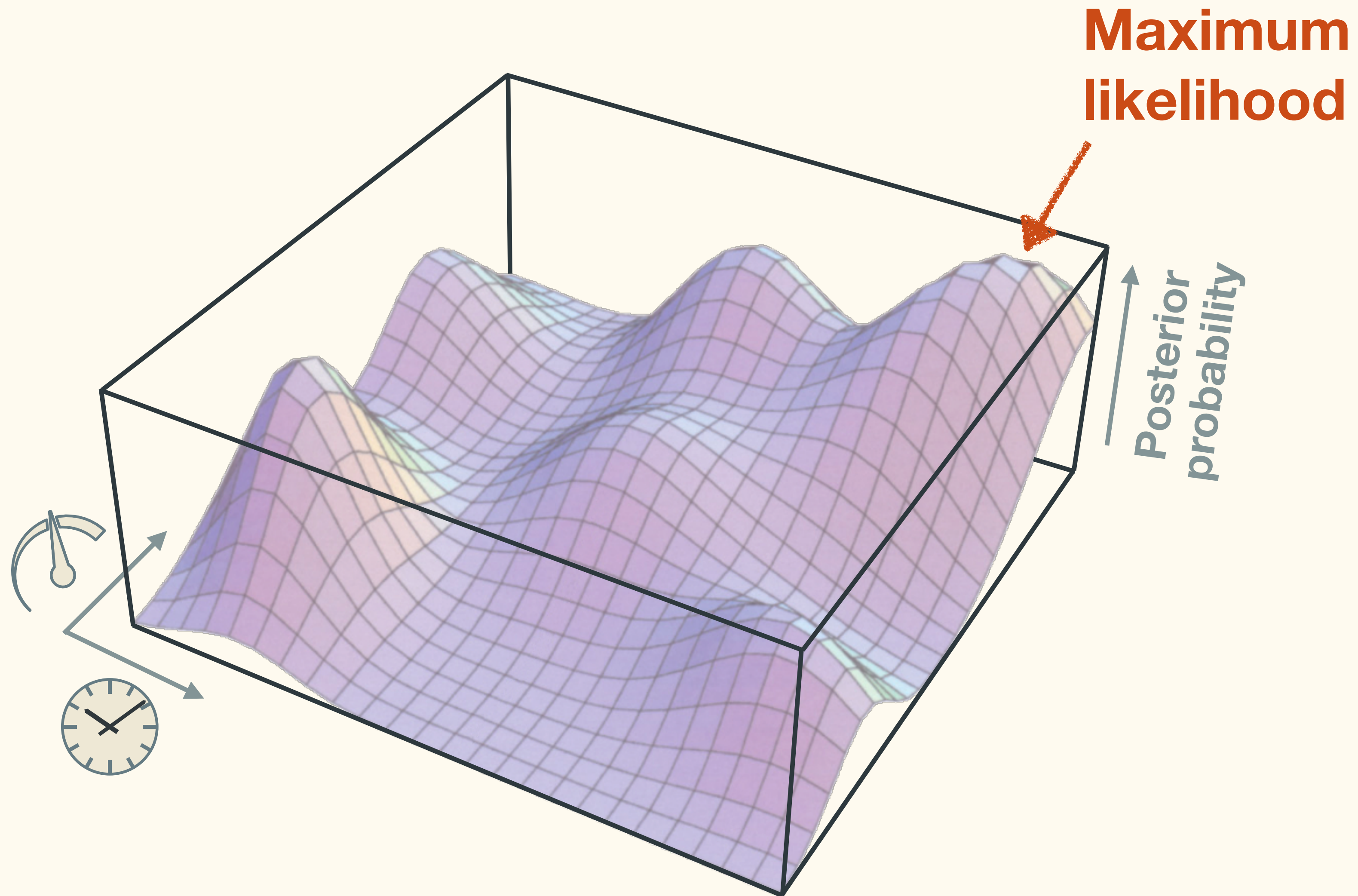
MCMC

Markov-chain Monte Carlo



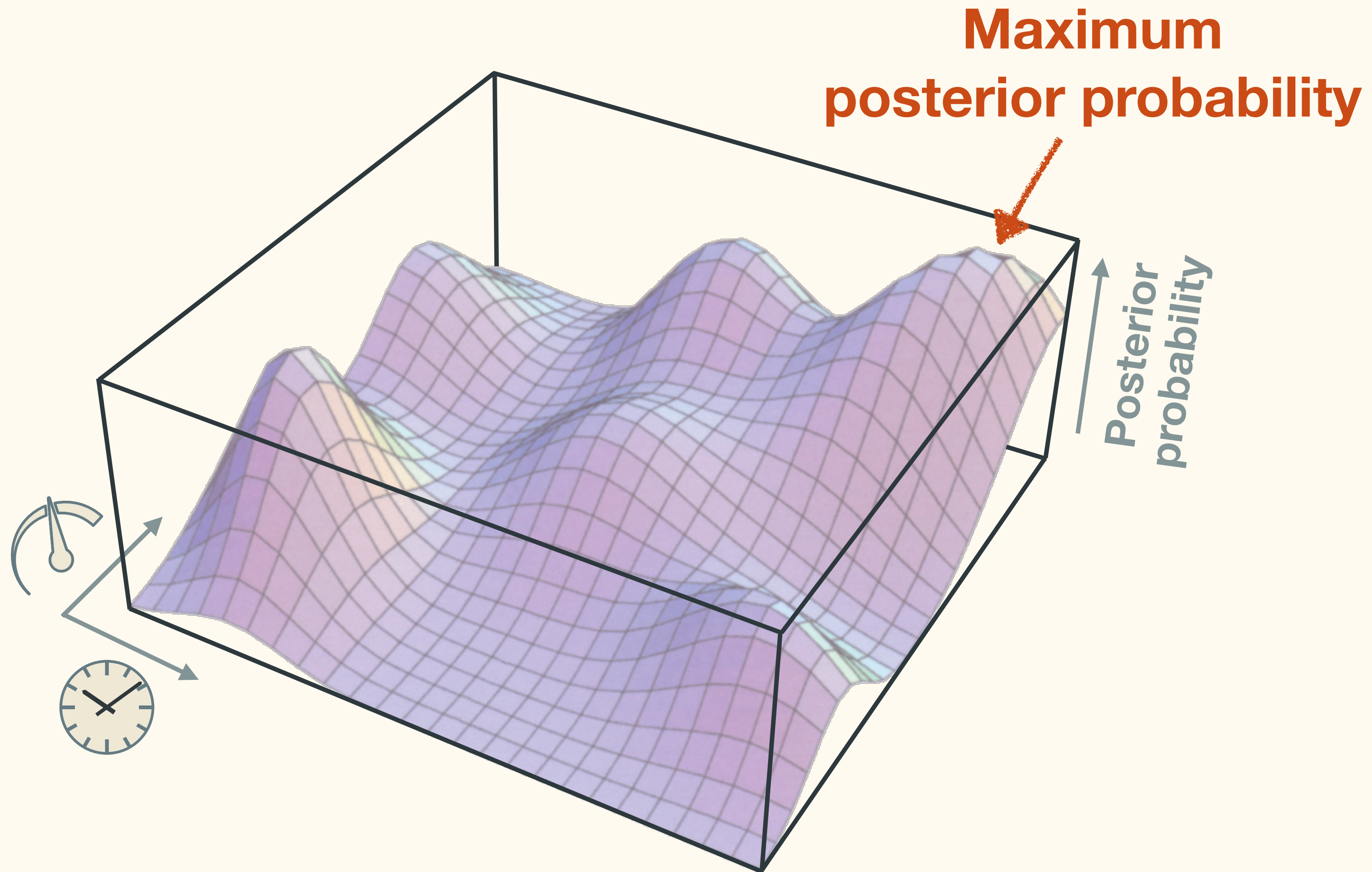
MCMC

Markov-chain Monte Carlo



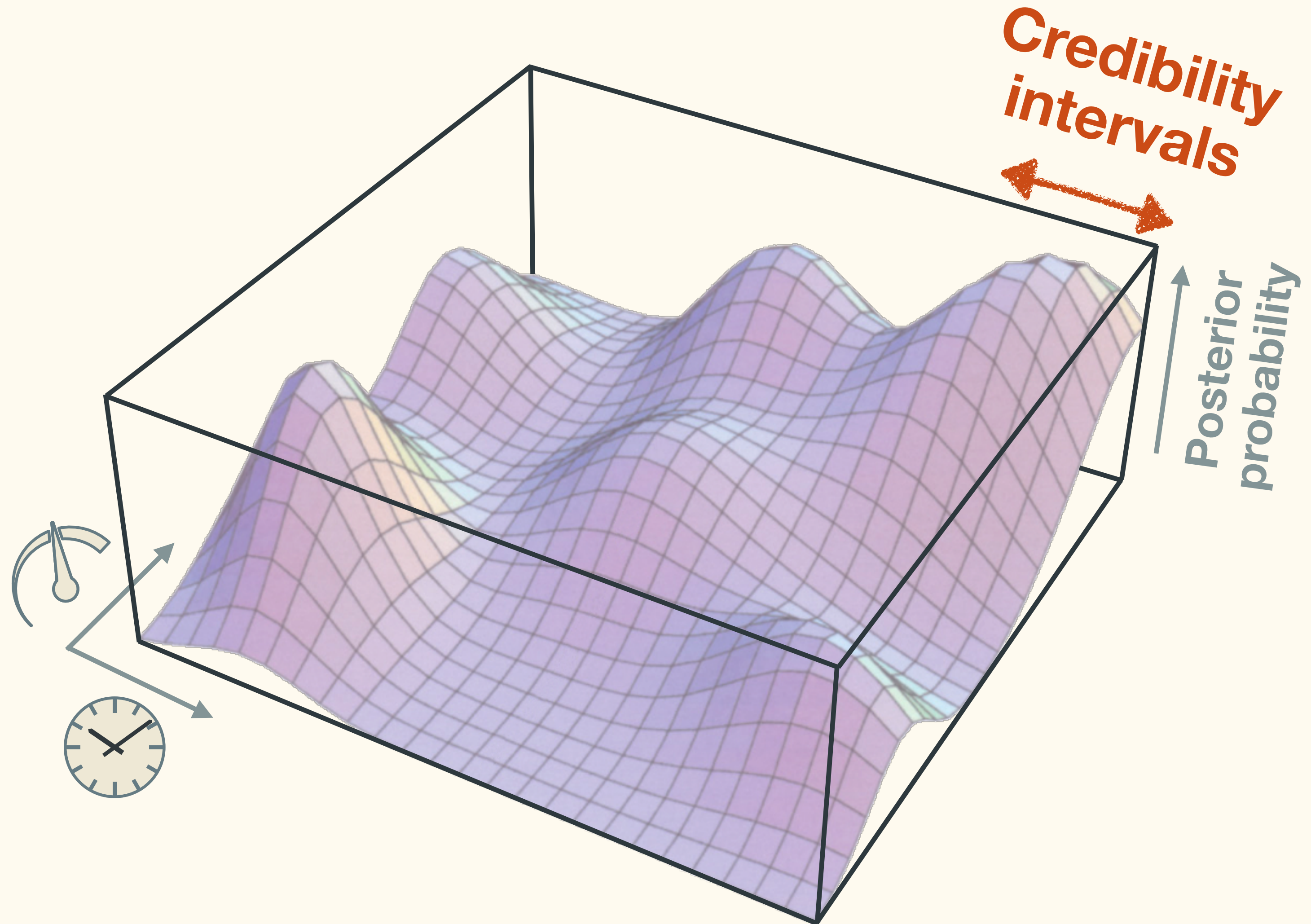
MCMC

Markov-chain Monte Carlo



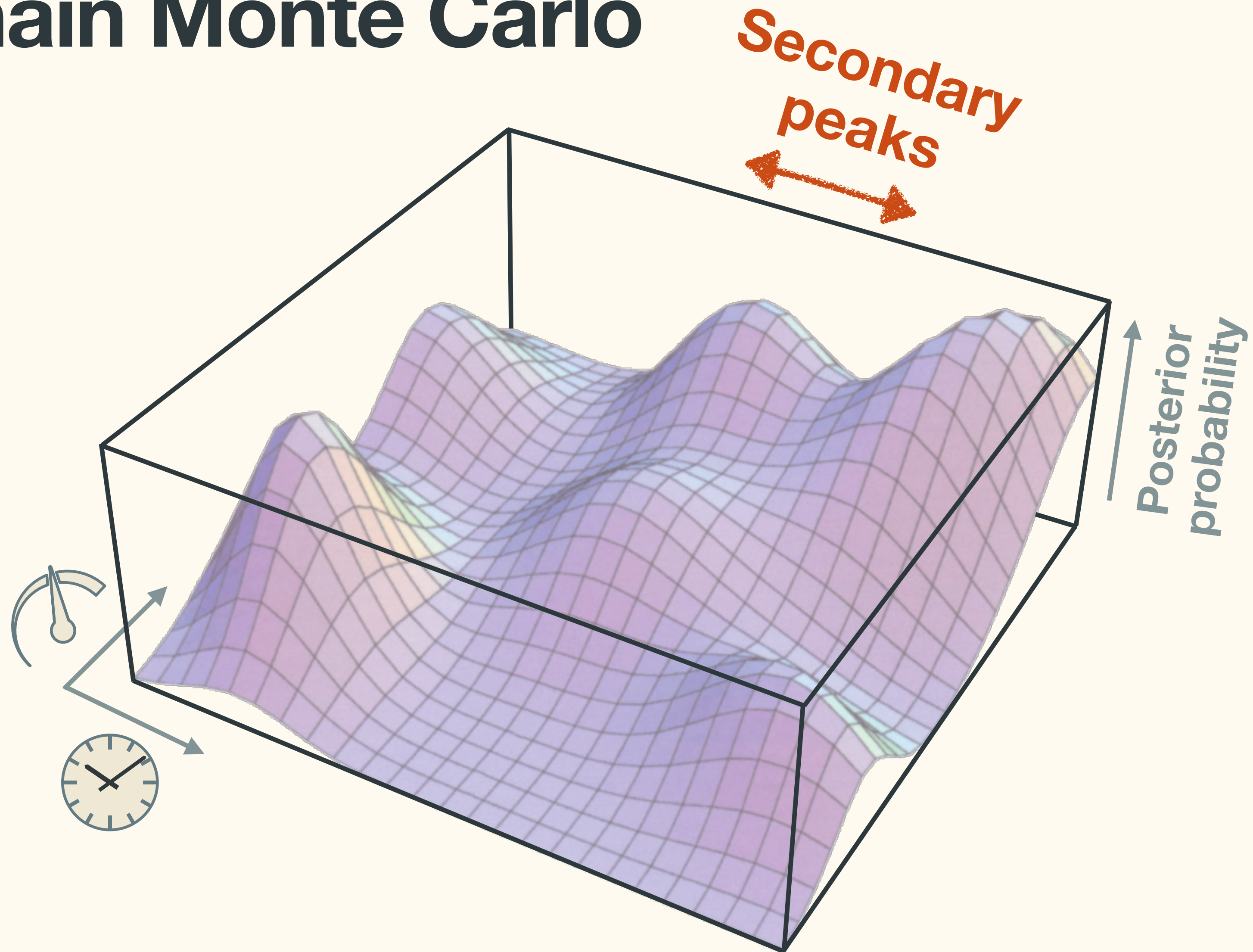
MCMC

Markov-chain Monte Carlo



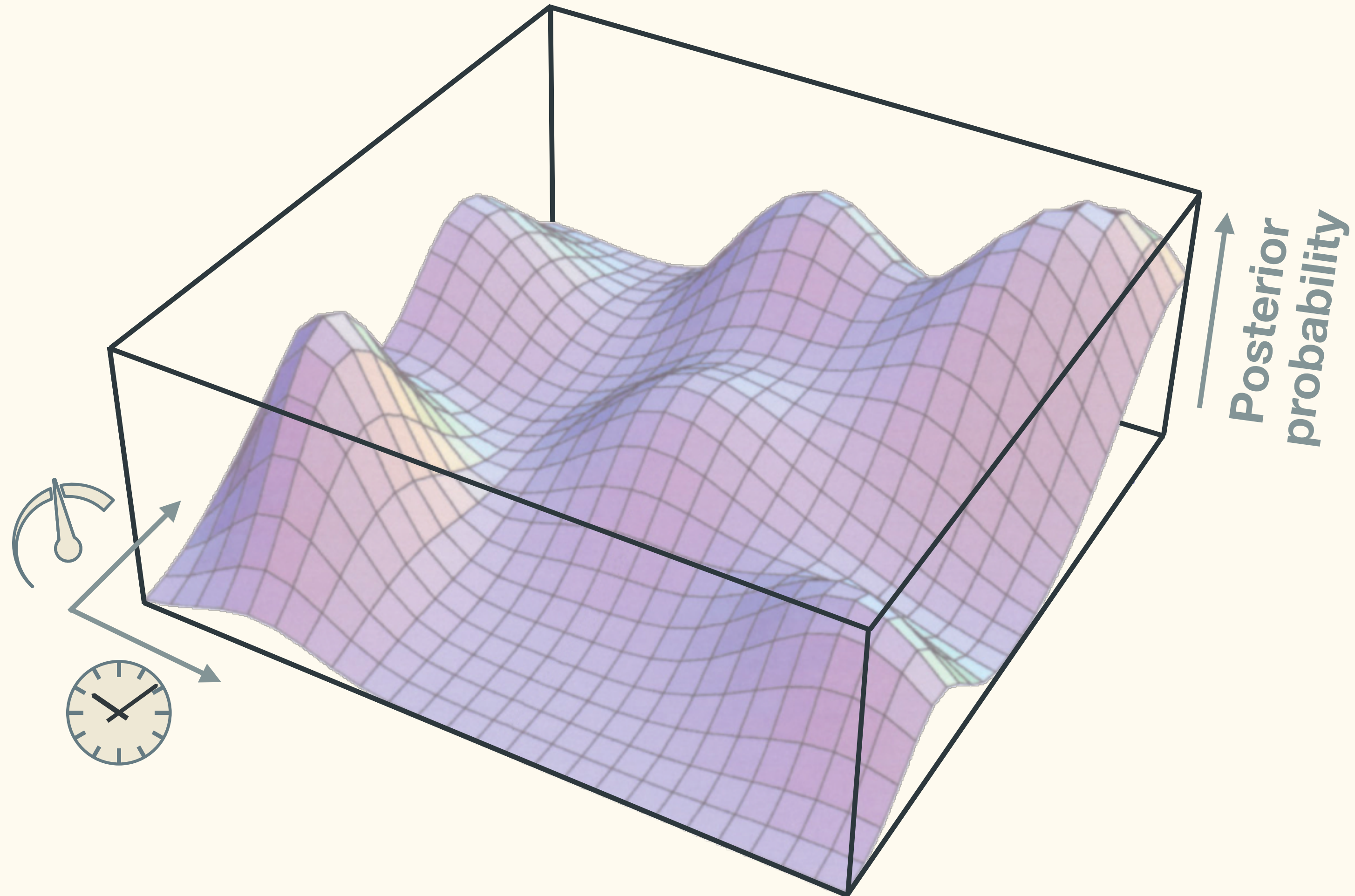
MCMC

Markov-chain Monte Carlo



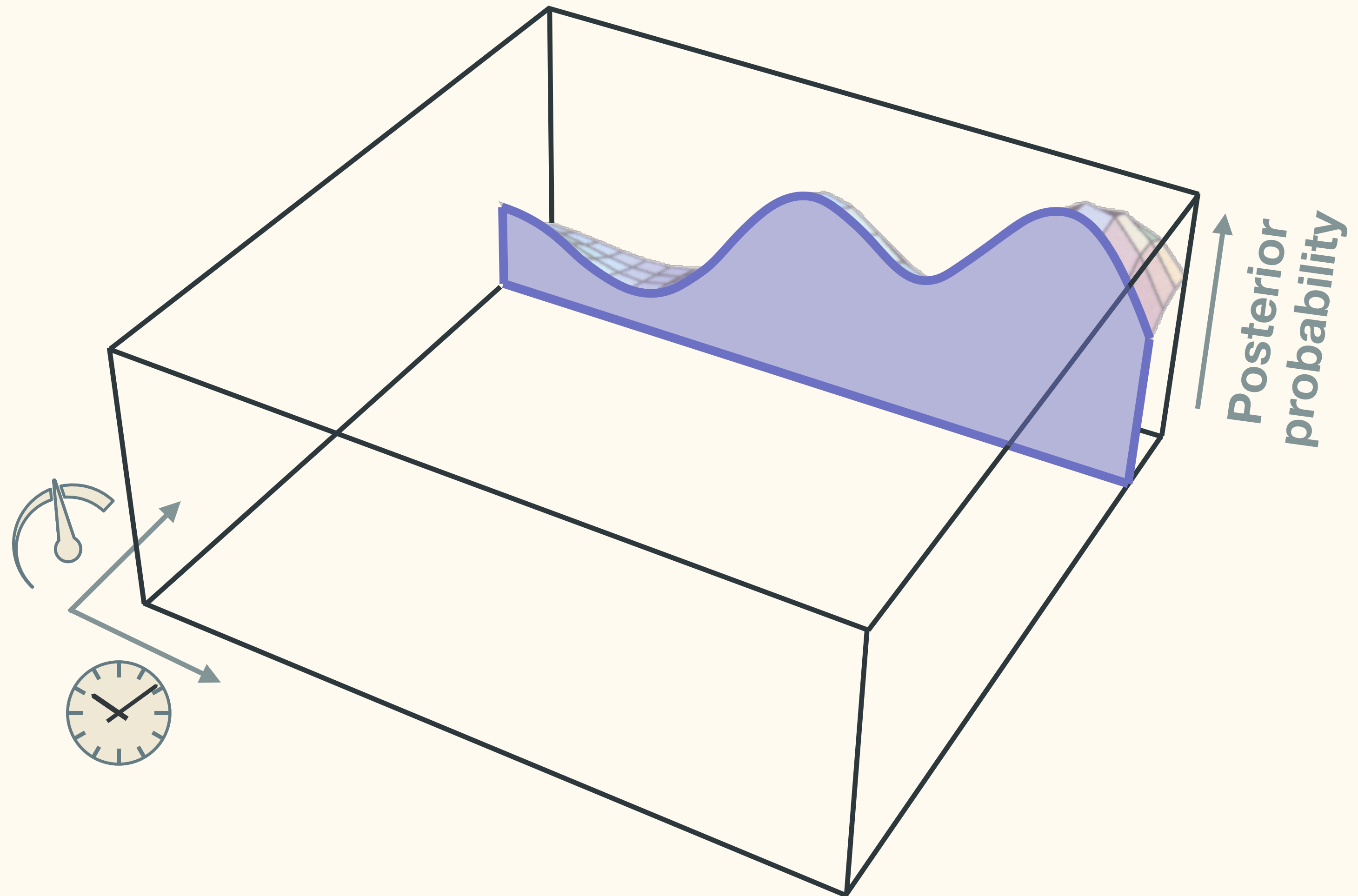
MCMC

Markov-chain Monte Carlo



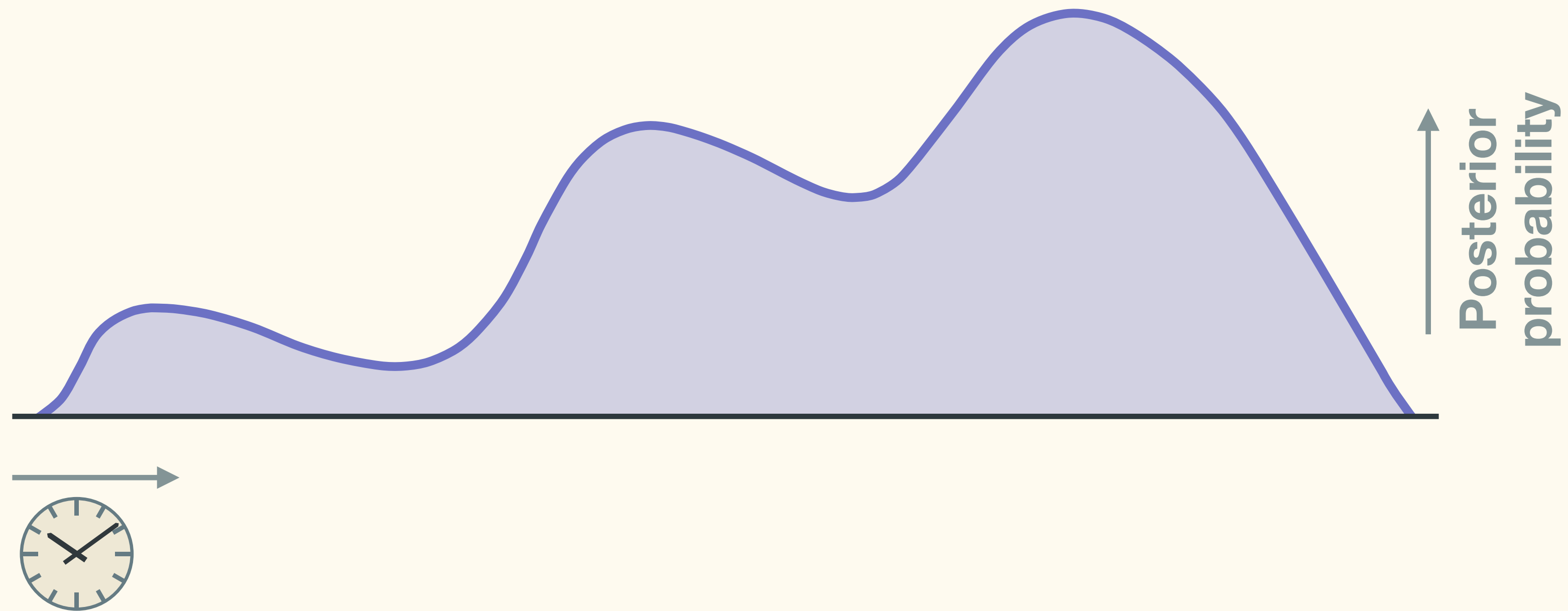
MCMC

Markov-chain Monte Carlo



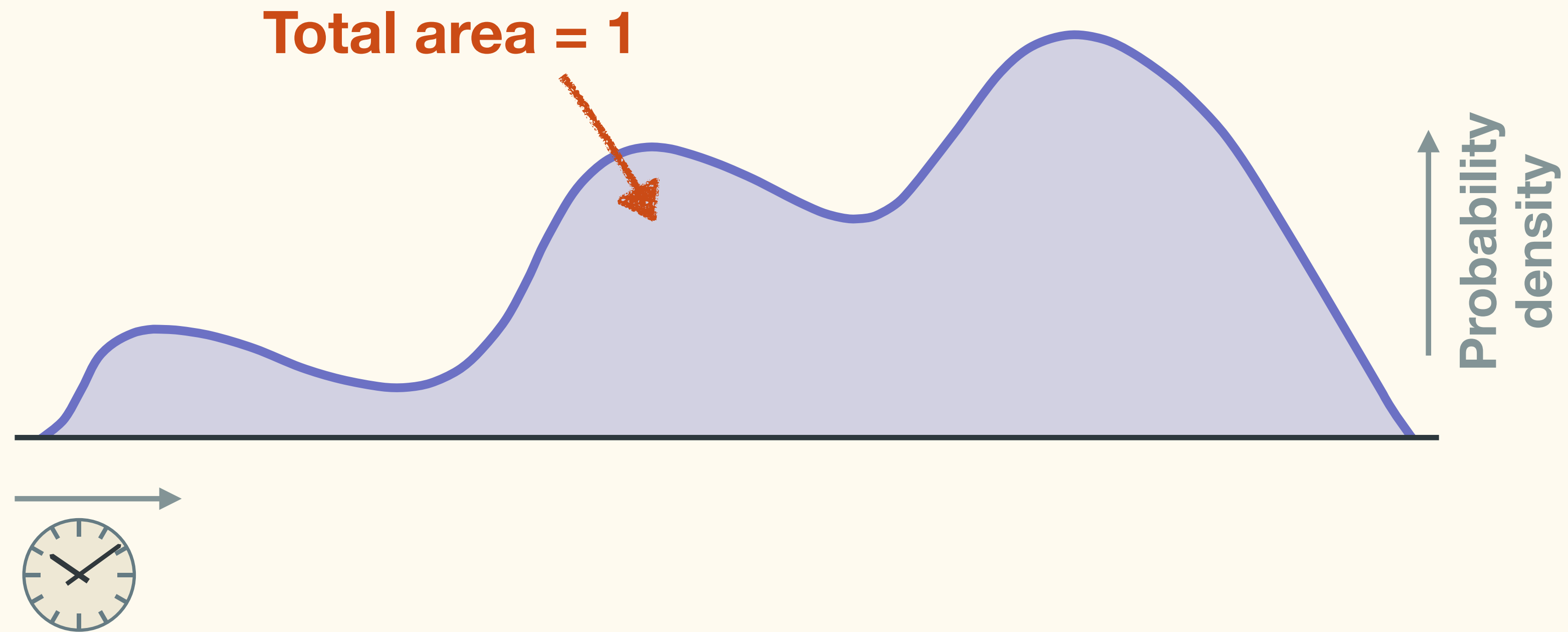
MCMC

Markov-chain Monte Carlo



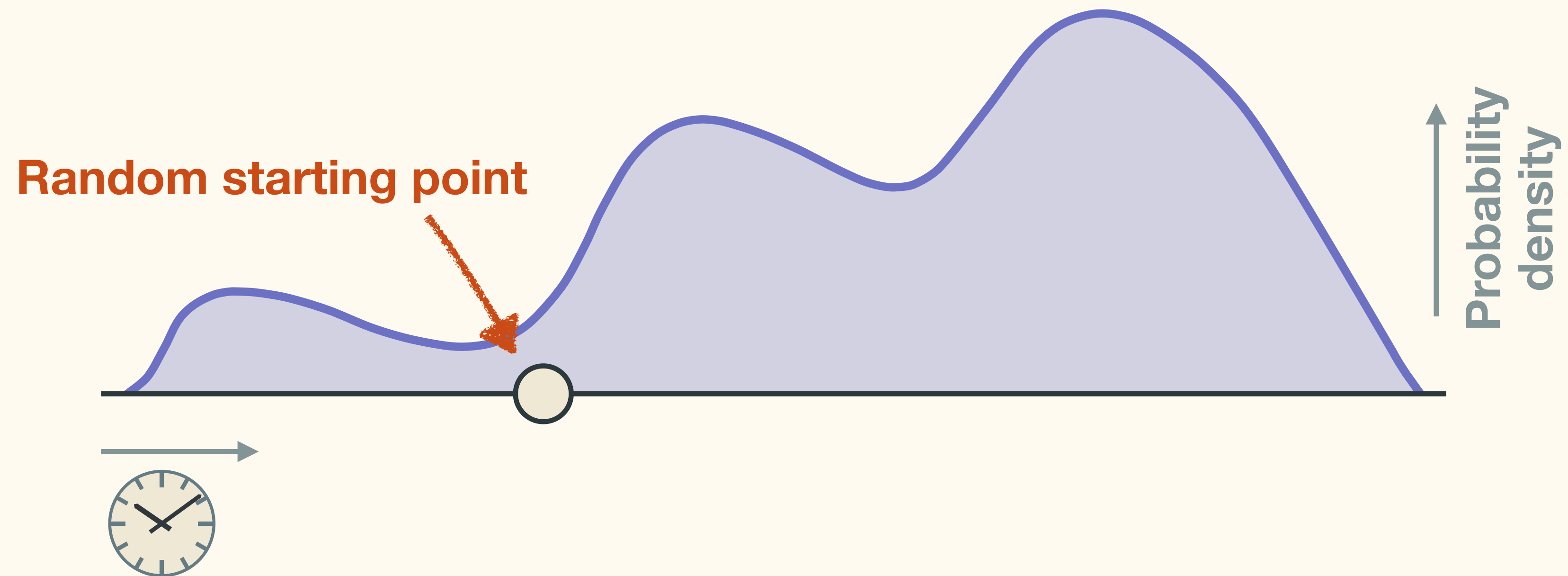
MCMC

Markov-chain Monte Carlo



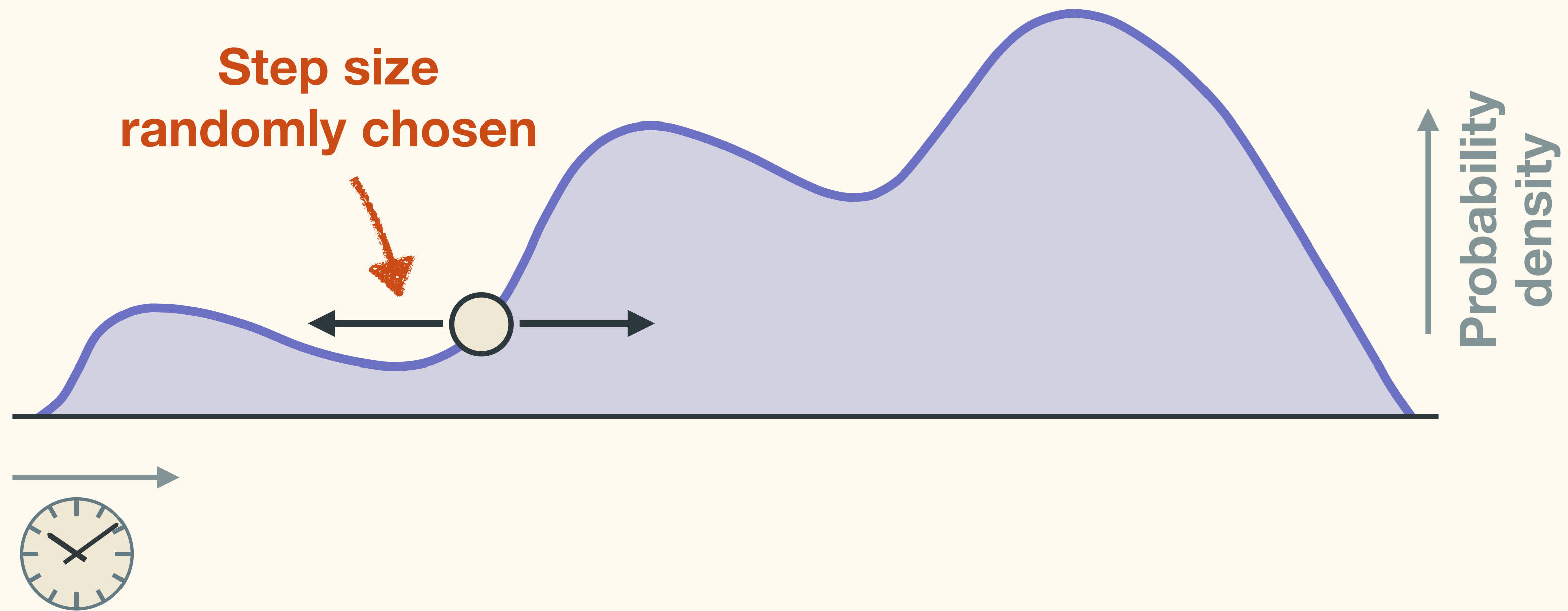
MCMC

Markov-chain Monte Carlo



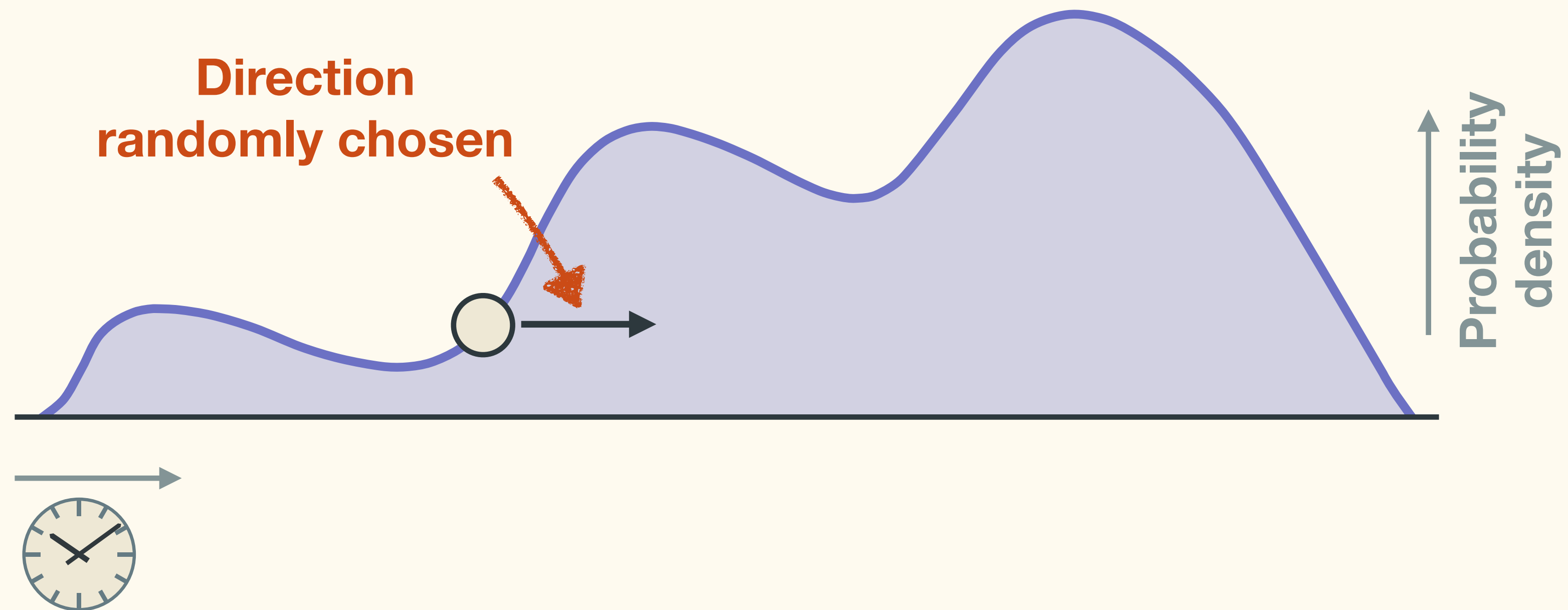
MCMC

Markov-chain Monte Carlo



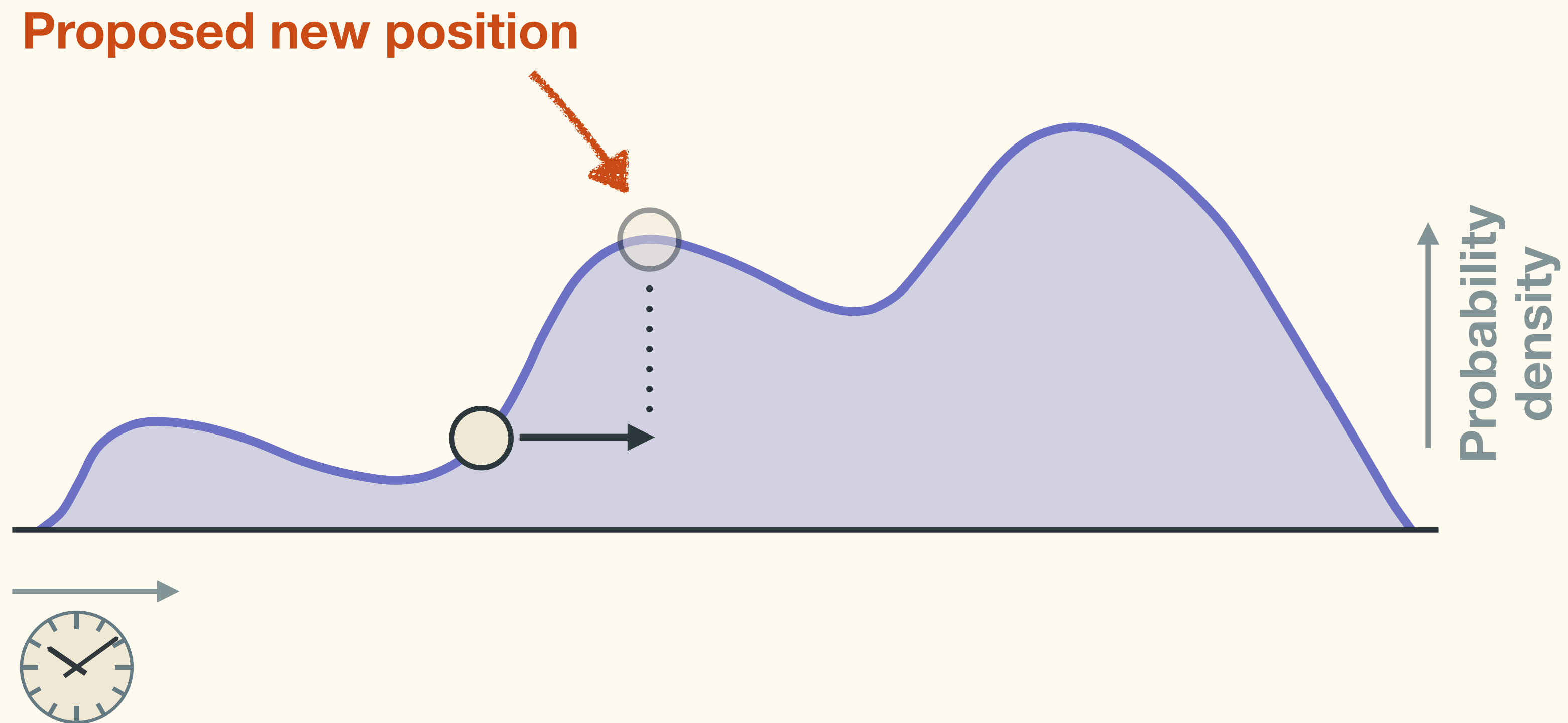
MCMC

Markov-chain Monte Carlo



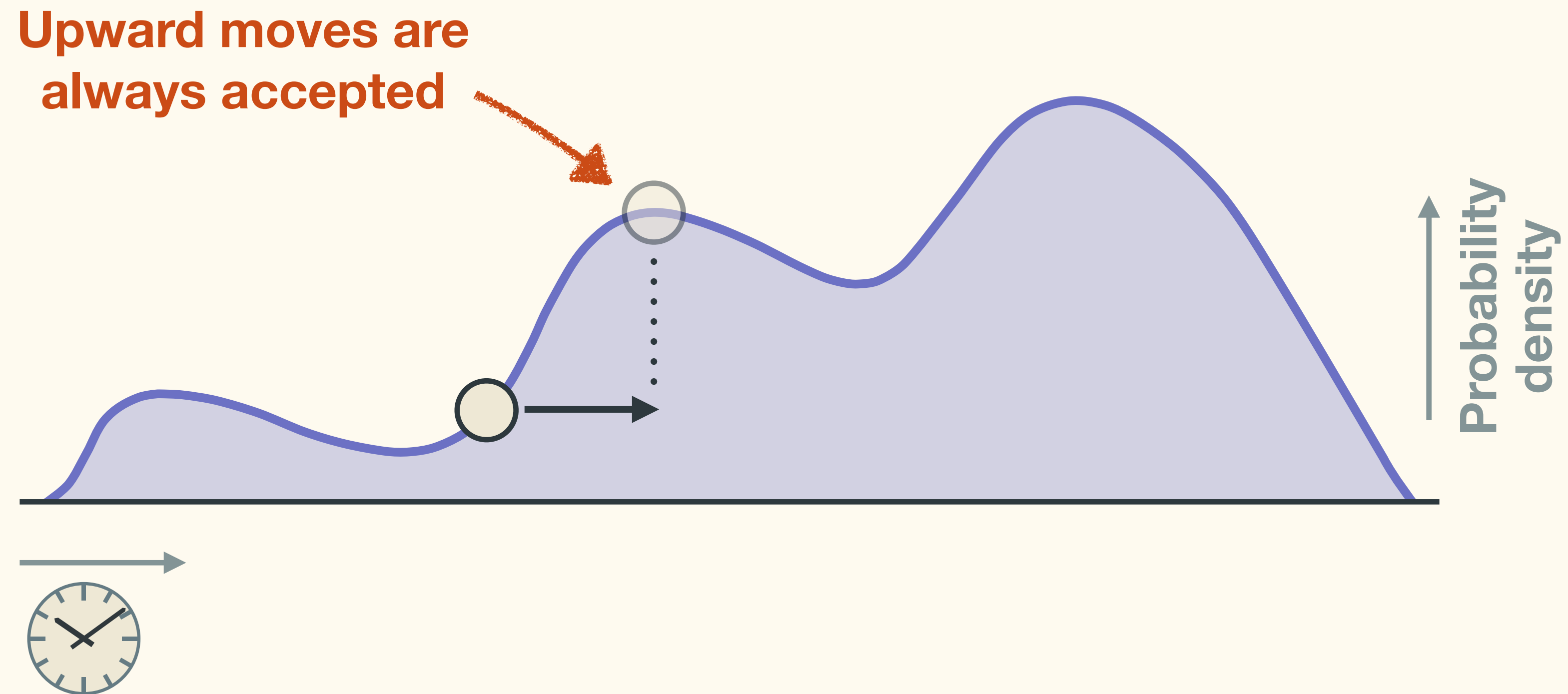
MCMC

Markov-chain Monte Carlo



MCMC

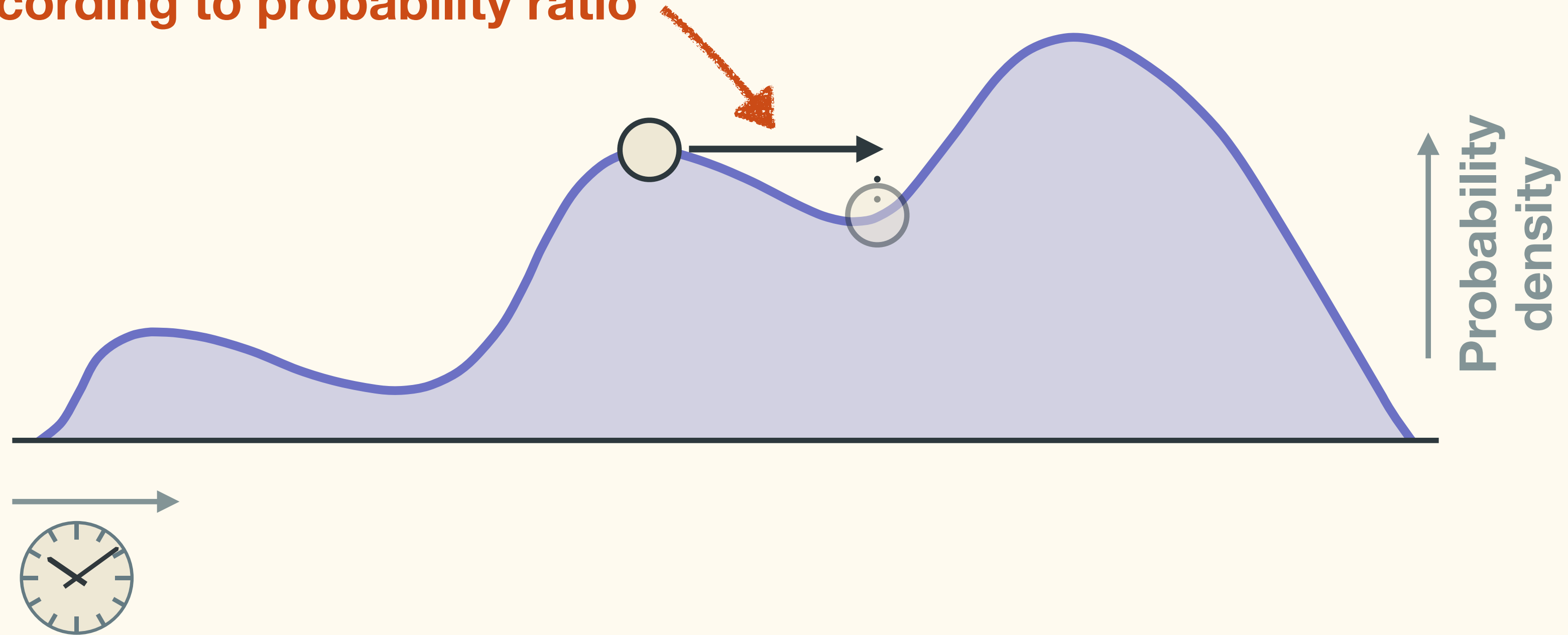
Markov-chain Monte Carlo



MCMC

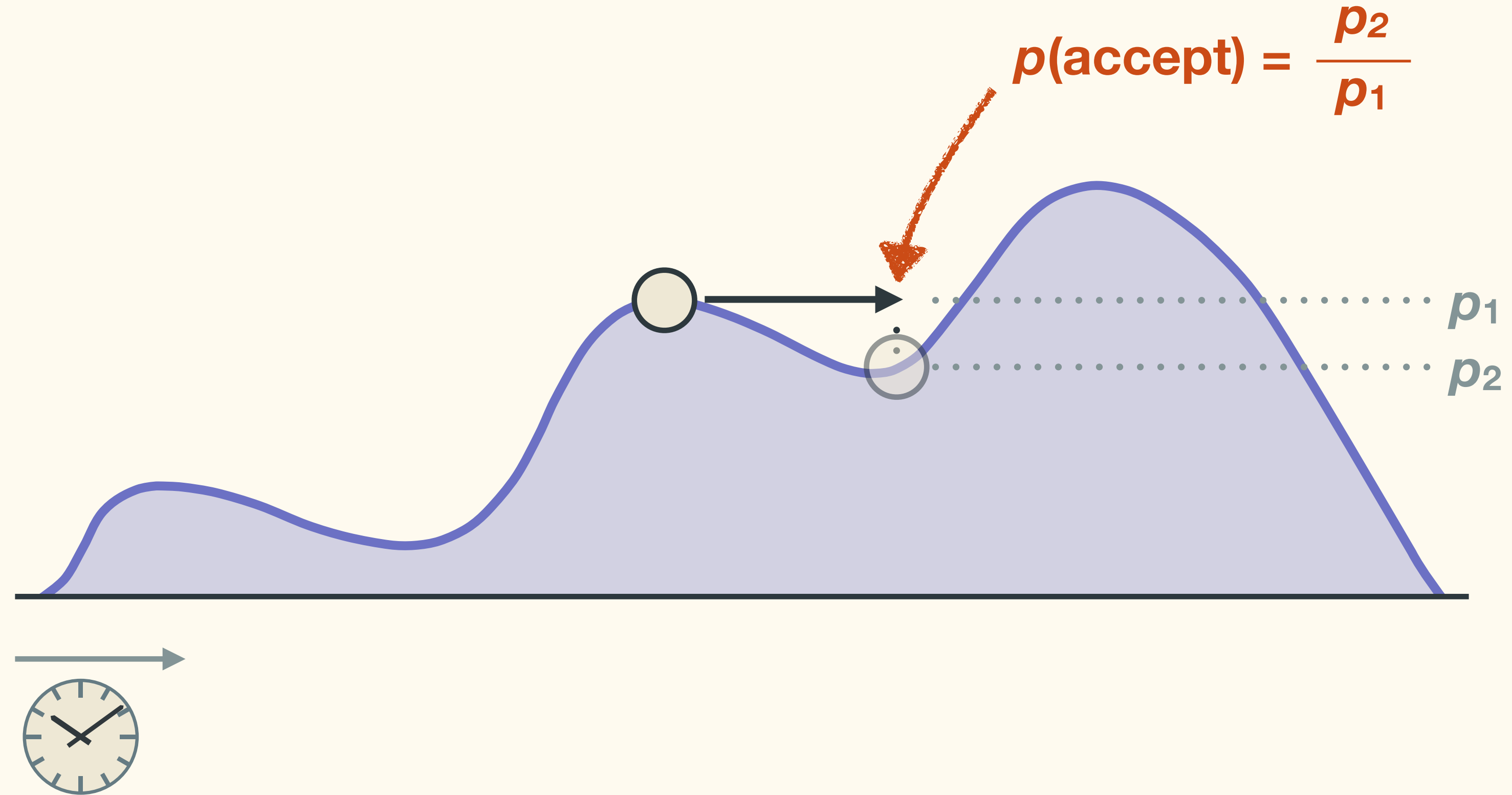
Markov-chain Monte Carlo

Downward moves are accepted according to probability ratio



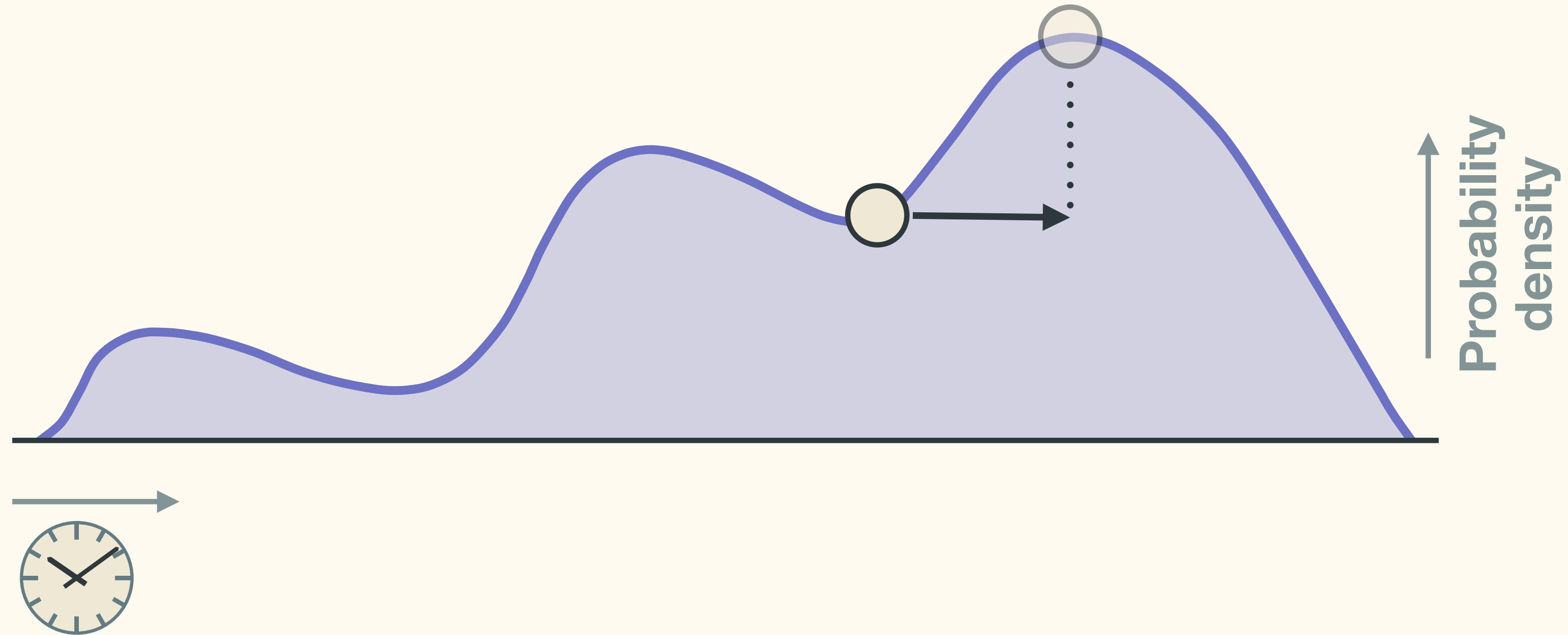
MCMC

Markov-chain Monte Carlo



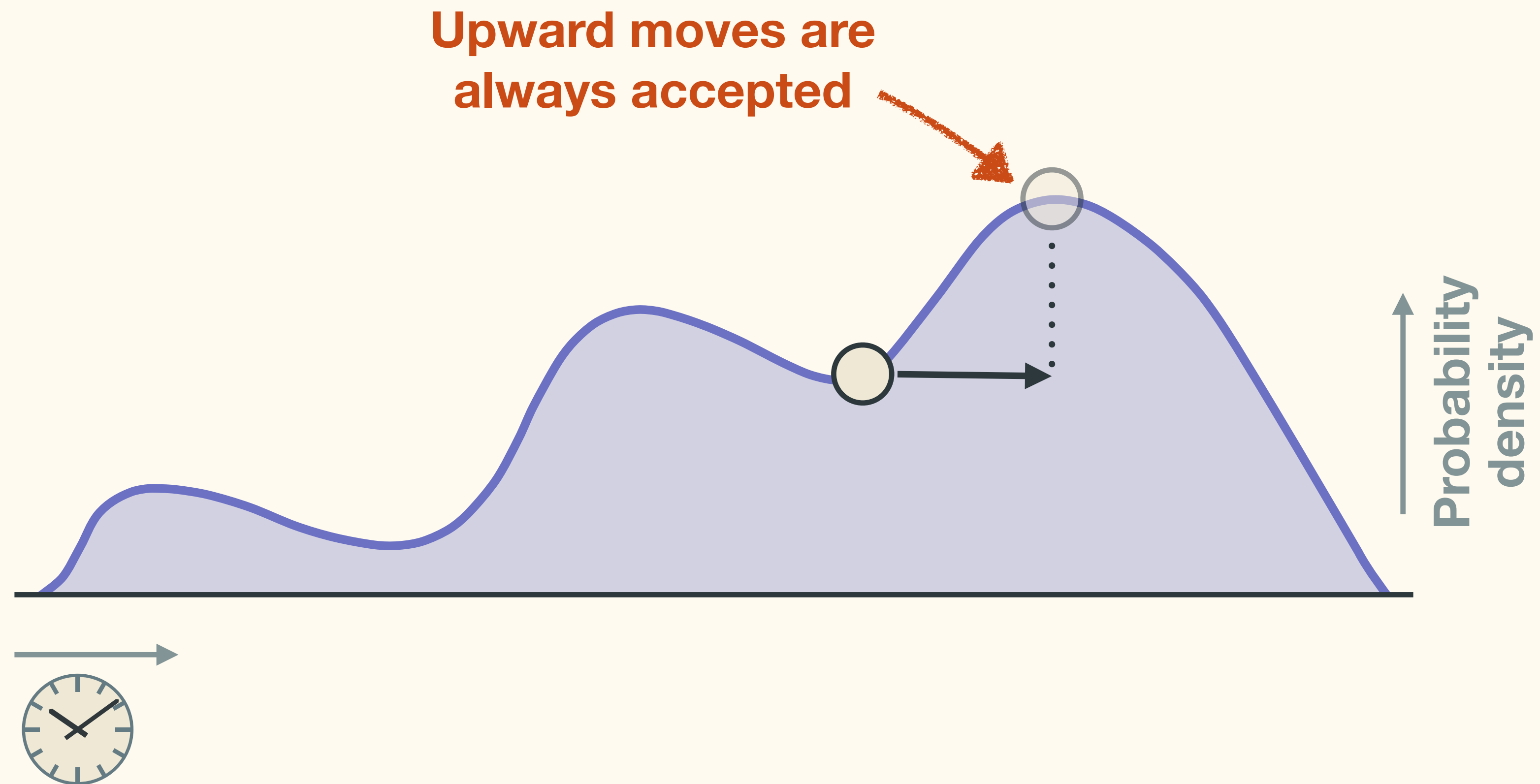
MCMC

Markov-chain Monte Carlo



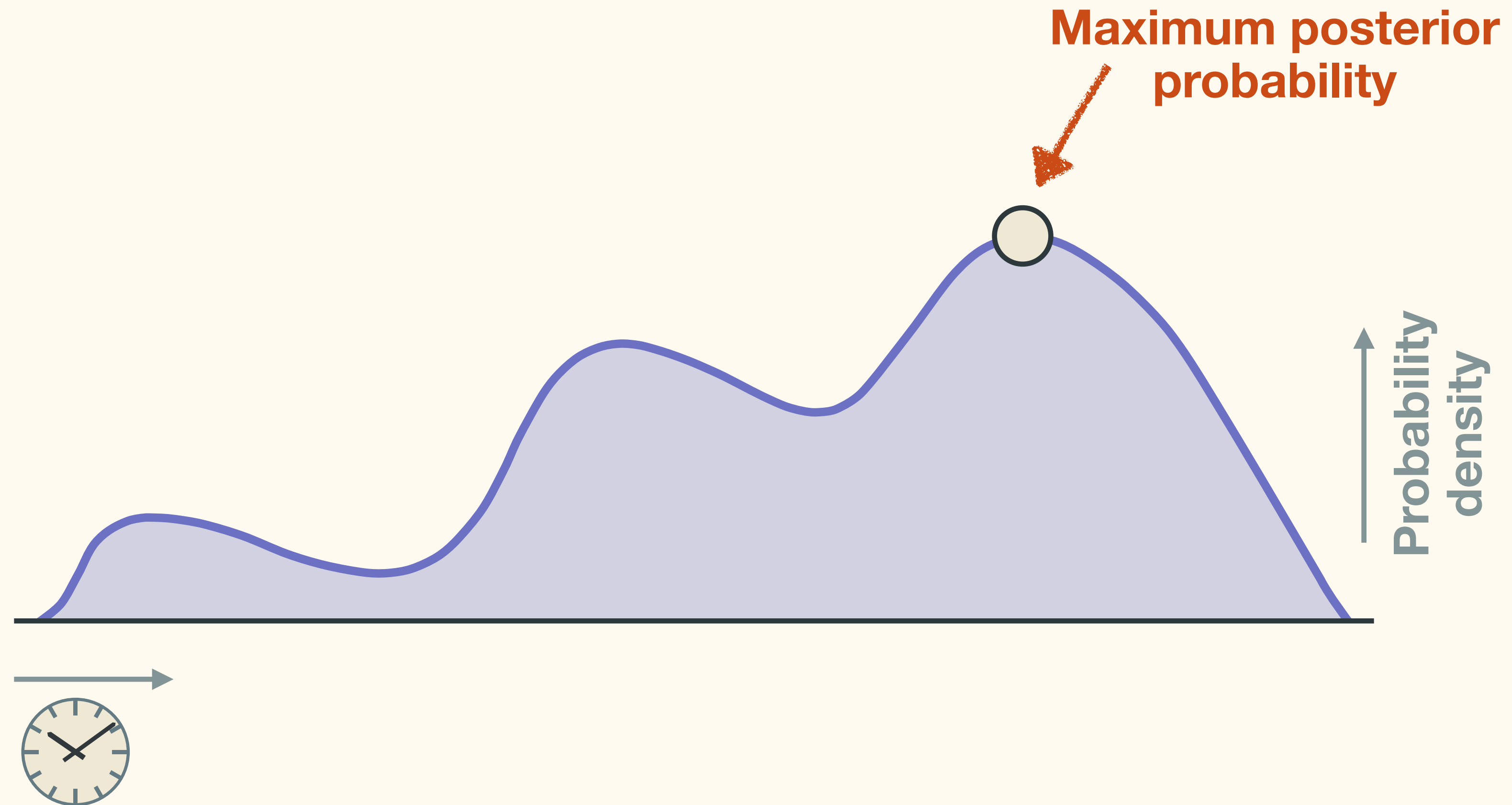
MCMC

Markov-chain Monte Carlo



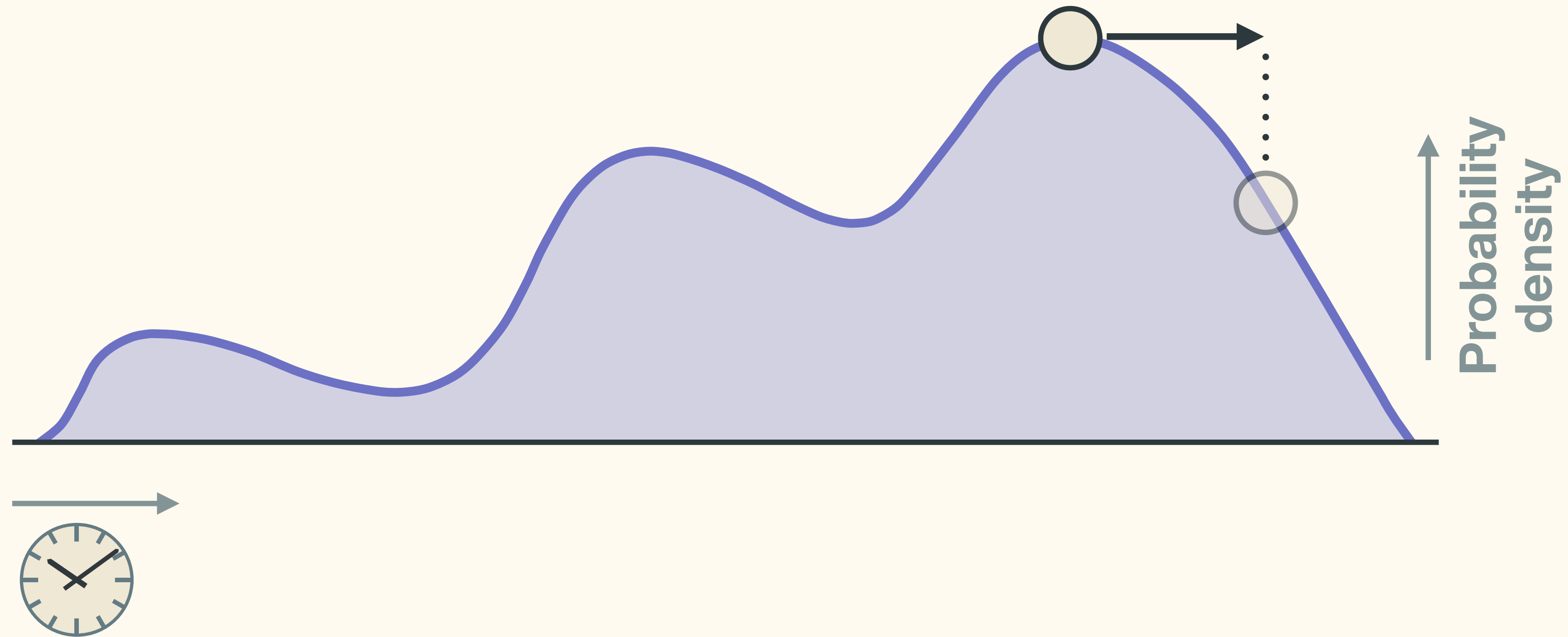
MCMC

Markov-chain Monte Carlo



MCMC

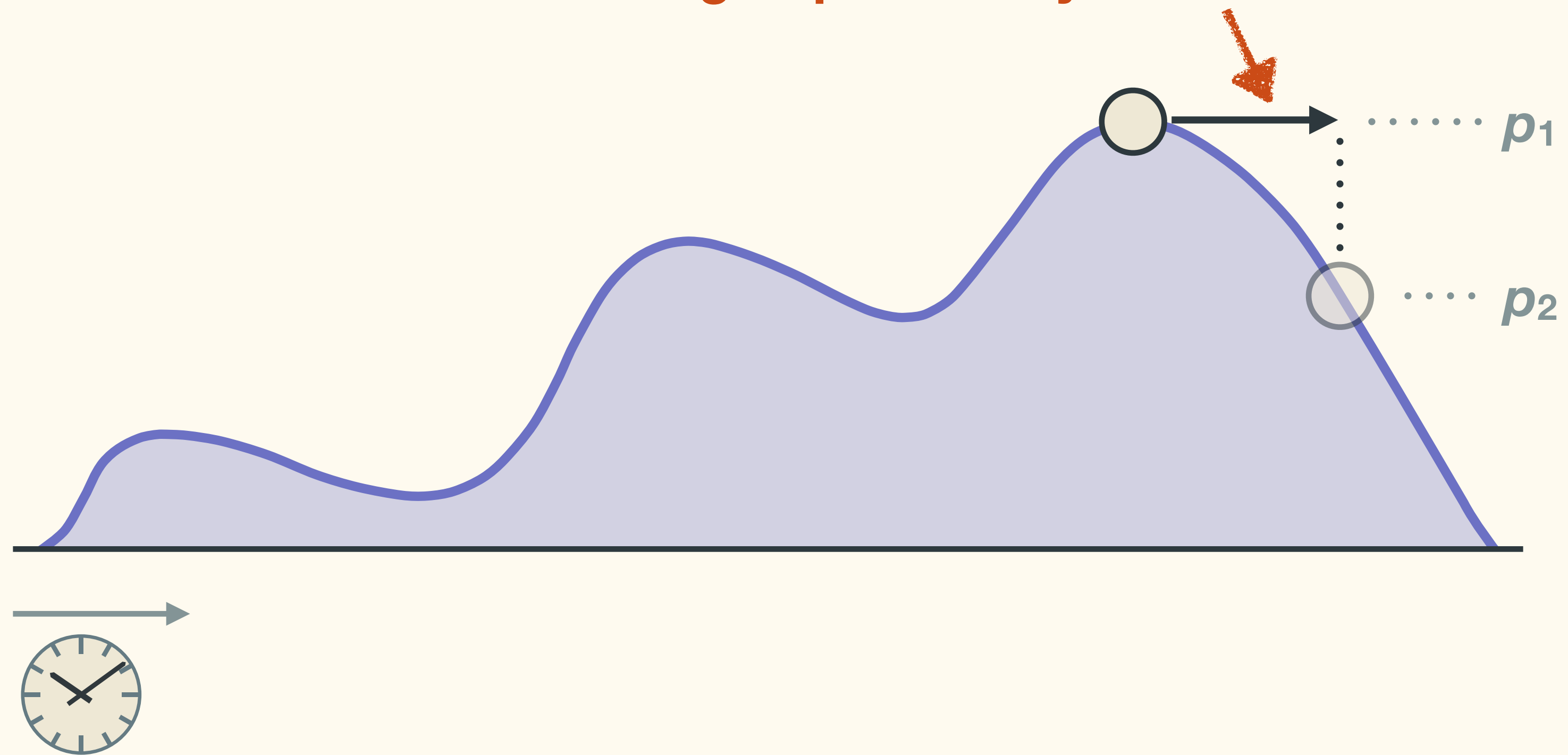
Markov-chain Monte Carlo



MCMC

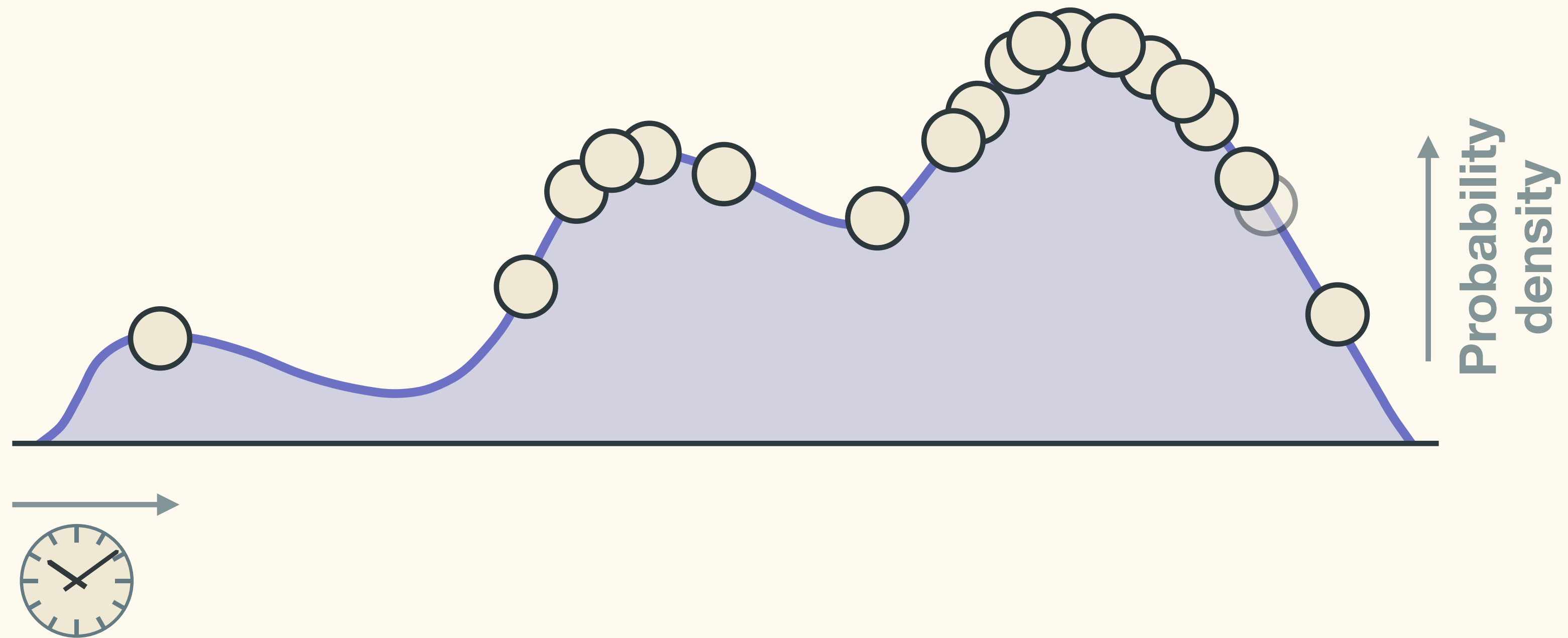
Markov-chain Monte Carlo

Downward moves are accepted according to probability ratio



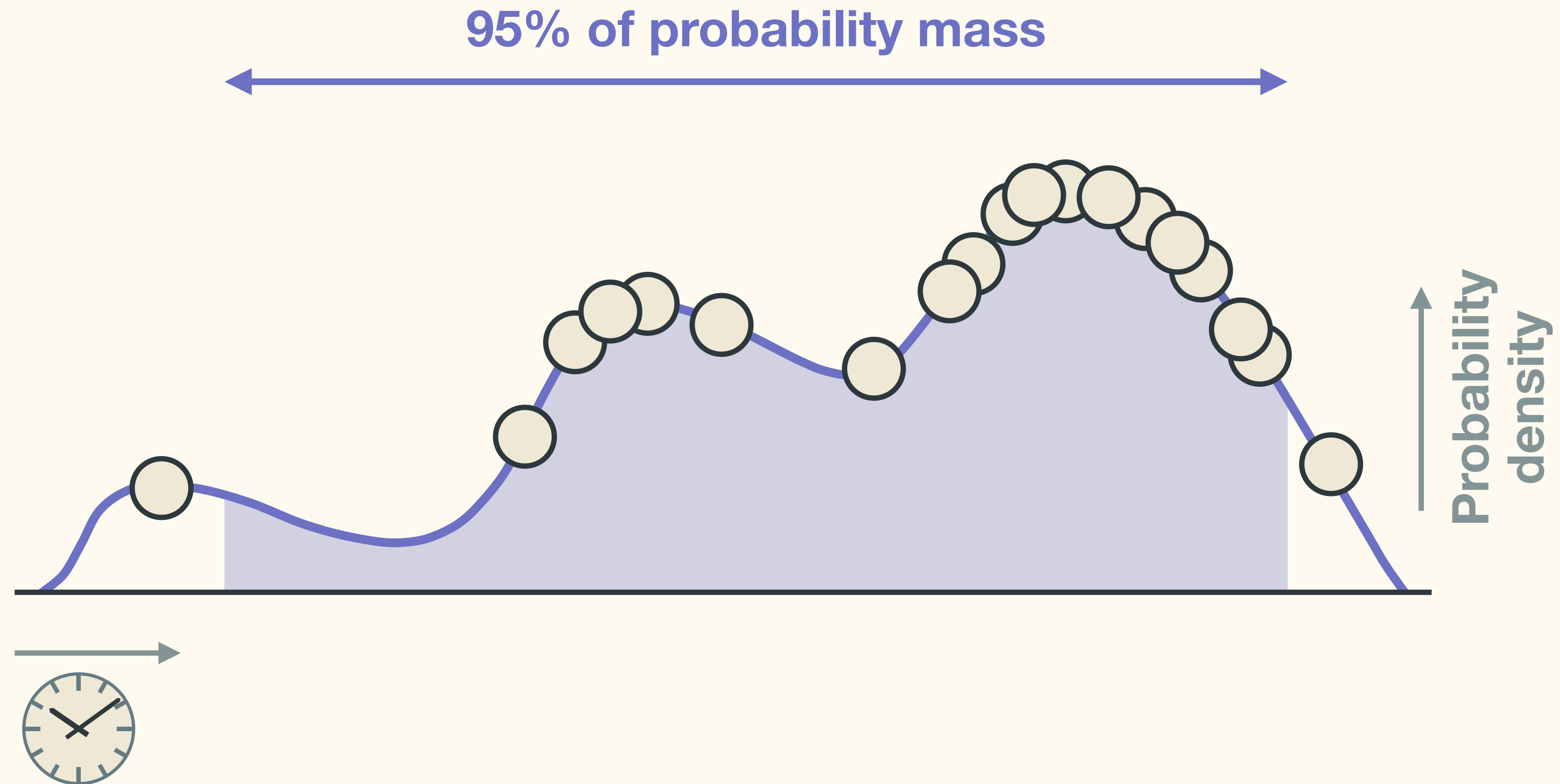
MCMC

Markov-chain Monte Carlo



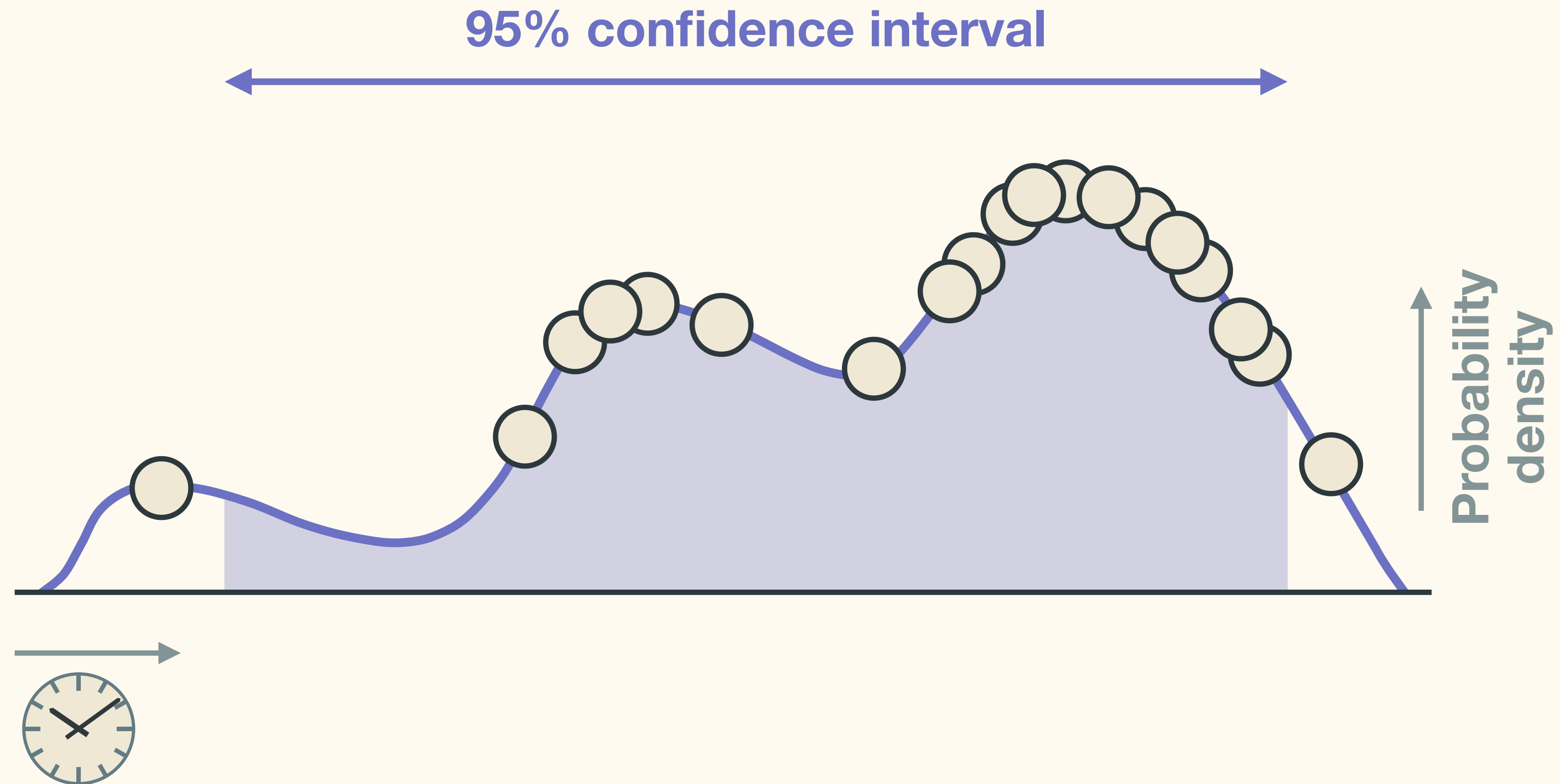
MCMC

Markov-chain Monte Carlo



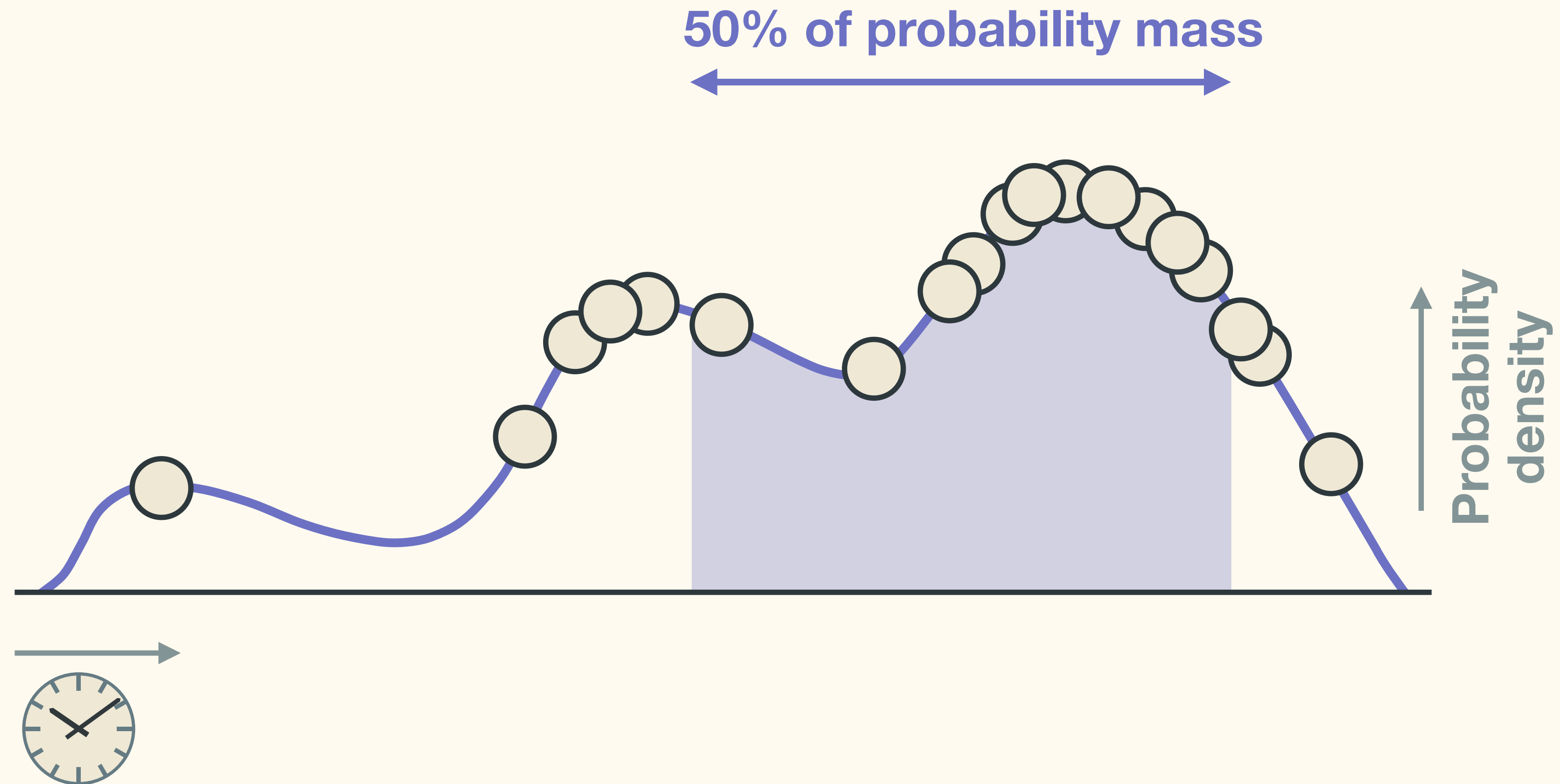
MCMC

Markov-chain Monte Carlo



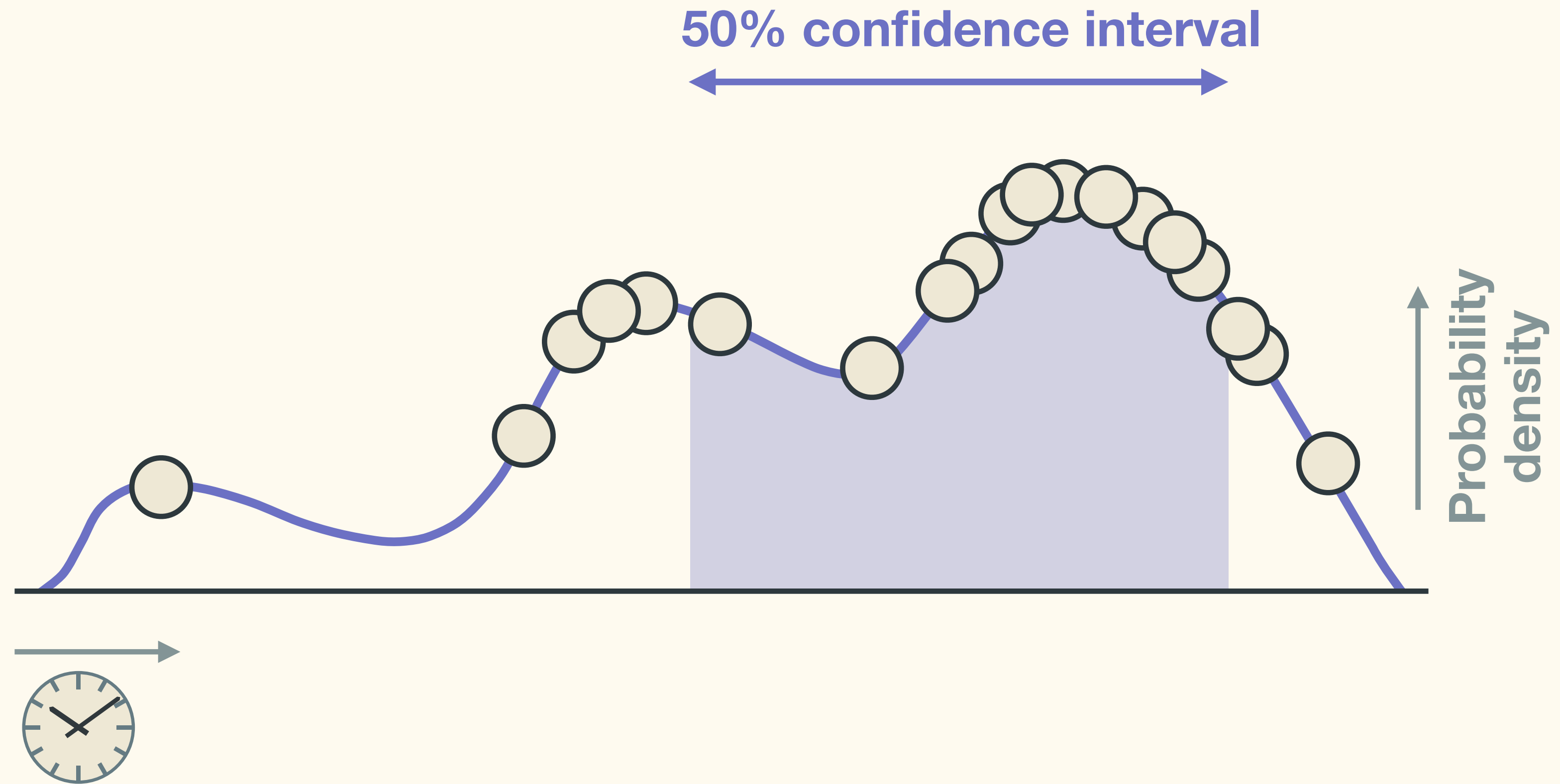
MCMC

Markov-chain Monte Carlo



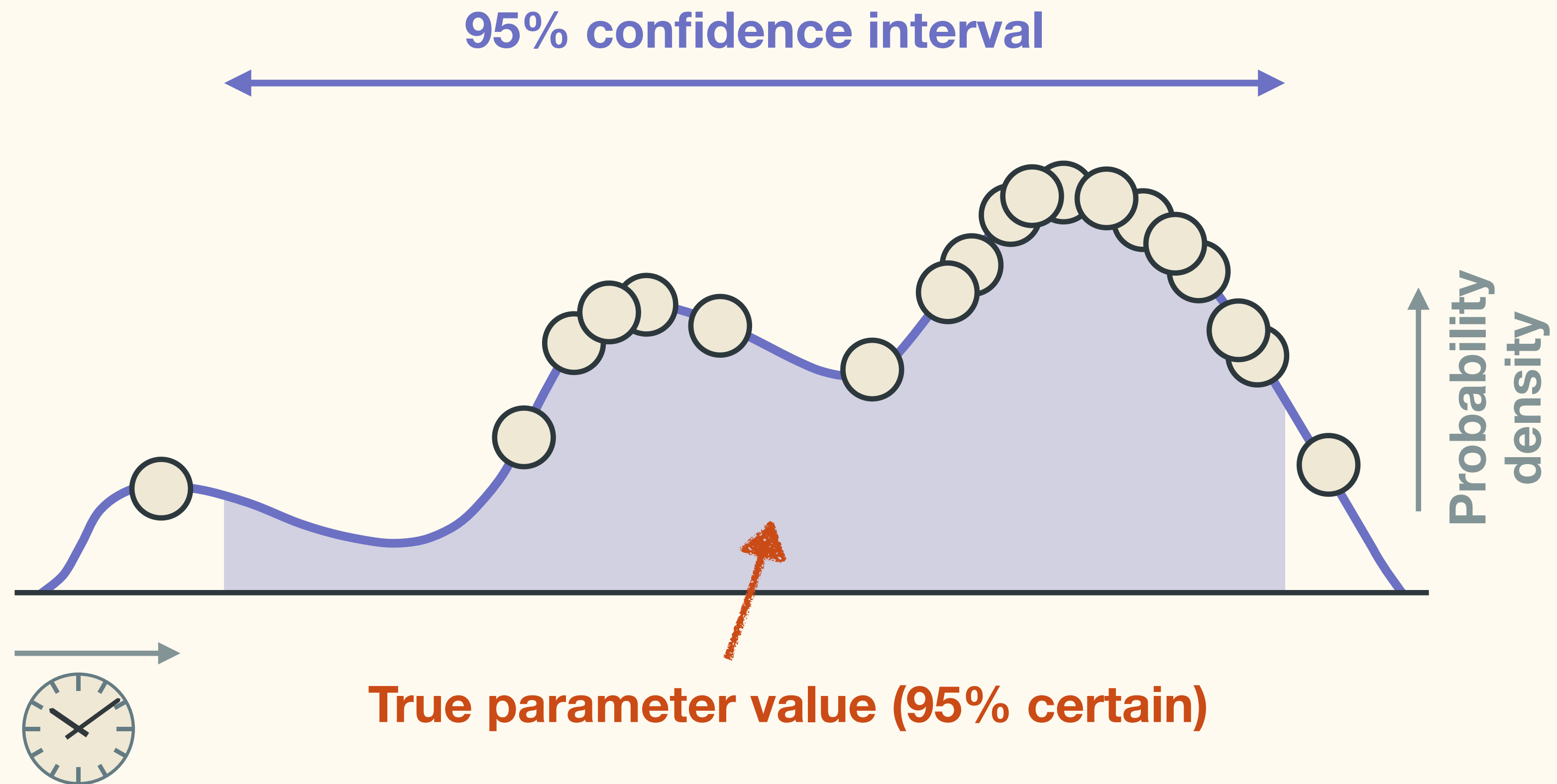
MCMC

Markov-chain Monte Carlo



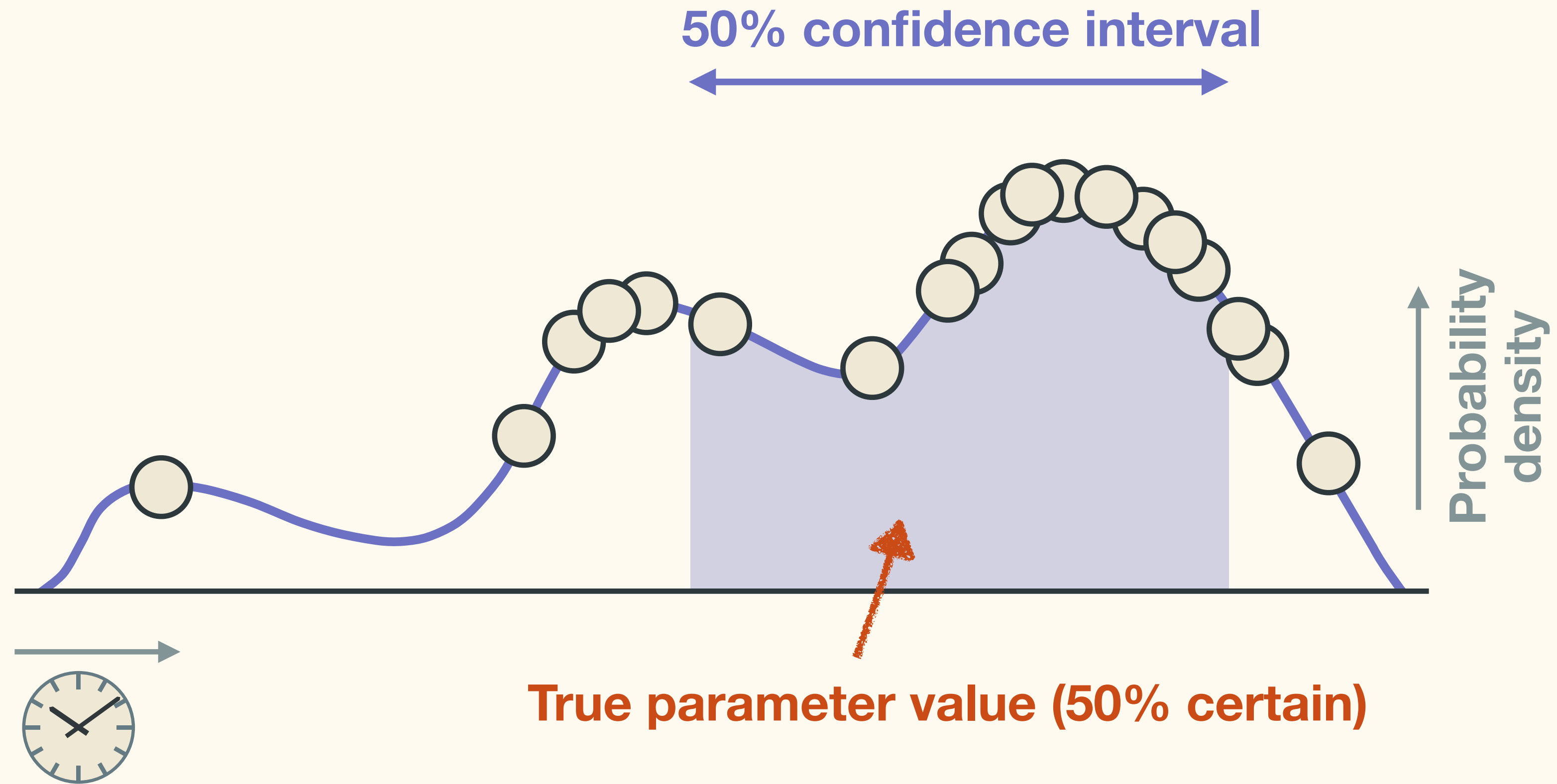
MCMC

Markov-chain Monte Carlo



MCMC

Markov-chain Monte Carlo



MCMC Robot

Thanks

Thanks

$$P\left(\begin{array}{c} \text{I'M NEAR} \\ \text{THE OCEAN} \end{array} \middle| \begin{array}{c} \text{I PICKED UP} \\ \text{A SEASHELL} \end{array}\right) =$$

$$\frac{P\left(\begin{array}{c} \text{I PICKED UP} \\ \text{A SEASHELL} \end{array} \middle| \begin{array}{c} \text{I'M NEAR} \\ \text{THE OCEAN} \end{array}\right) P\left(\begin{array}{c} \text{I'M NEAR} \\ \text{THE OCEAN} \end{array}\right)}{P\left(\begin{array}{c} \text{I PICKED UP} \\ \text{A SEASHELL} \end{array}\right)}$$

$$P\left(\begin{array}{c} \text{I PICKED UP} \\ \text{A SEASHELL} \end{array}\right)$$



CRASHHHH
SPLOOSH

STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND DON'T HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.