Today

- Finish survey of JC69 relatives
- Probability fundamentals
- How to calculate a likelihood:
  - 1 sequence (not that useful)
  - 2 sequences (allows estimation of an evolutionary distance)
  - 4 sequences (allows estimation of a tree)

"Transition/transversion ratio" vs. "transition/transversion *rate* ratio"

Possible transitions: Possible transversions:

$$\frac{E[\text{No. transitions}]}{E[\text{No. transversions}]} = \underline{\hspace{6cm}} =$$

F81 rate matrix

"To" state

	A	C	G	T
A				
"From" state **C**				
G				
T				

Felsenstein (1981)

parameters?

equivalence to JC69

HKY85 rate matrix

"To" state

		A	C	G	T
"From" state	A				
	C				
	G				
	T				

Hasegawa, Kishino, & Yano (1985)

parameters?

equivalence to JC69, F81

Evolutionary distances for several common models

Model	Expected no. substitutions: $v = \{r\}t$
JC69	$\nu = \{3\beta\}t$
F81	$\nu = \{2\beta\left(\pi_R\pi_Y + \pi_A\pi_G + \pi_C\pi_T\right)\}t$
K80	$\nu = \{\beta\left(\kappa + 2\right)\}t$
HKY85	$\nu = \{2\beta\left[\pi_R\pi_Y + \kappa\left(\pi_A\pi_G + \pi_C\pi_T\right)\right]\}t$

GTR rate matrix

"To" state

	A	C	G	T
A				
C				
G				
T				

"From" state

Lanave et al. (1984)
Tavaré (1986)

parameters?

equivalence to JC69, F81, K80, HKY

Why do we need the term *likelihood*?

Outcome	Fair coin model	Two-heads model	
H	0.5	1	
T	0.5	0	Likelihoods of models given one particular data outcome are not expected to sum to 1.0
	1	1	

Probabilities of data outcomes given one particular model sum to 1.0

Probability of the **data** given the model

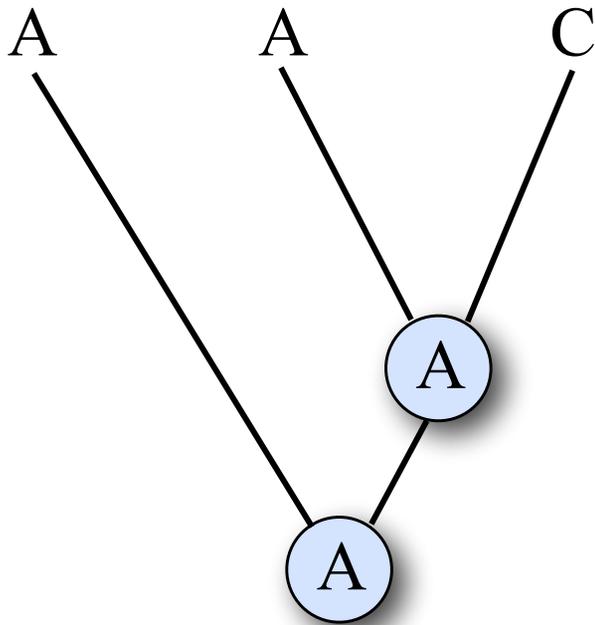**Likelihood** of the **model** given the data

Paul O. Lewis ~ Phylogenetics, Spring 2026

AND means MULTIPLY

Using 2 dice, what is the probability of

 AND ?

- *Multiply* probabilities if the component events must happen **simultaneously** (i.e. where you would naturally use the word AND when describing the problem)
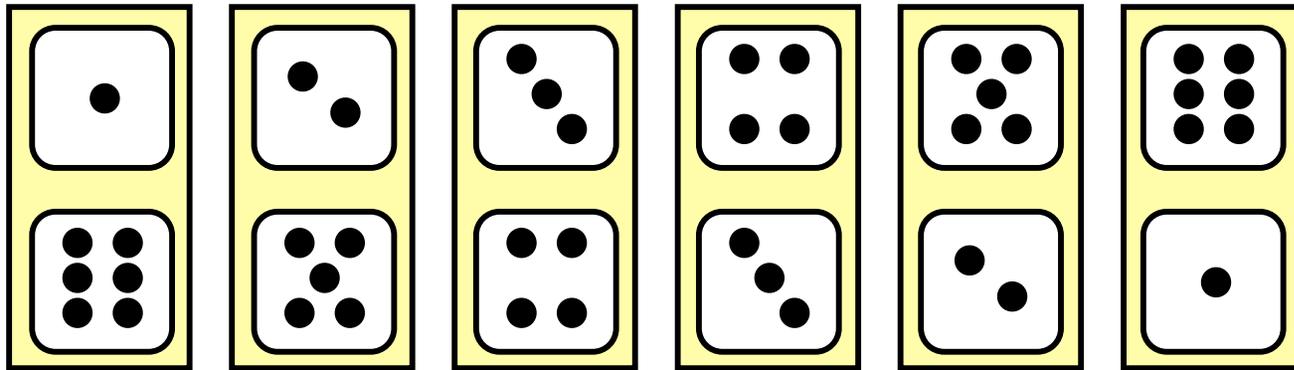
AND rule in phylogenetics

OR means ADD
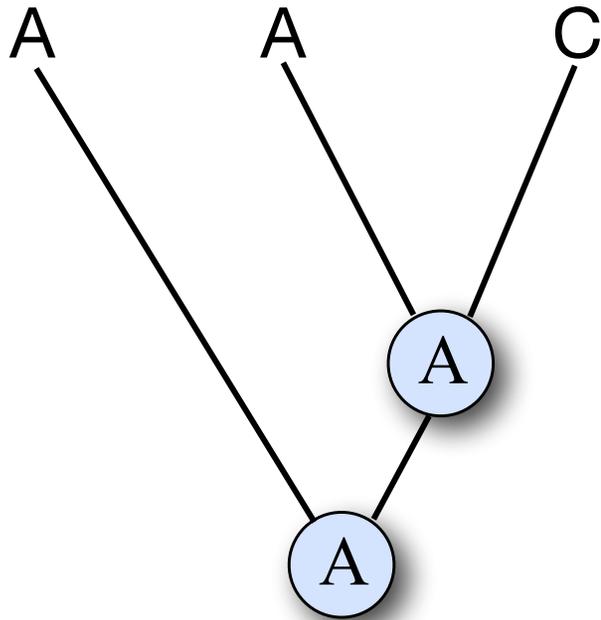
Using one die, what is the probability of

 OR ?

- *Add* probabilities if the component events are **mutually exclusive** (i.e. where you would naturally use the word OR in describing the problem)

Combining AND and OR

What is the probability that the sum of two dice is 7?

Using both AND and OR in phylogenetics

Likelihood of a single sequence

First 32 nucleotides of the ψη-globin gene of gorilla:

GAAGTCCTTGAGAAATAAACTGCACACACTGG

$$L =$$

$$\log L =$$