Brownian motion *and* phylogenetically independent contrasts

Liam J. Revell & Luke J. Harmon

12 December 2016

Brownian motion

- A stochastic model original developed to describe the movement of particles in a fluid.
- A model for the evolution of continuous traits.
- Traits change continuously through time.
- After some time the trait distribution follows a normal distribution

Outline

1. What is Brownian motion?
2. When might characters evolve by Brownian motion.
3. How can we simulate Brownian motion evolution on trees?

Brownian motion: The model

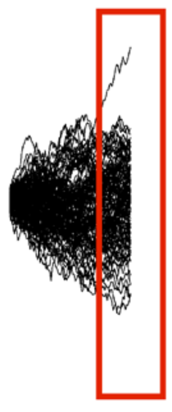
- Sometimes called a *Wiener process*.
- A continuous time stochastic process.
- Describes a ‘random-walk’ of evolution for a continuously valued character trait.

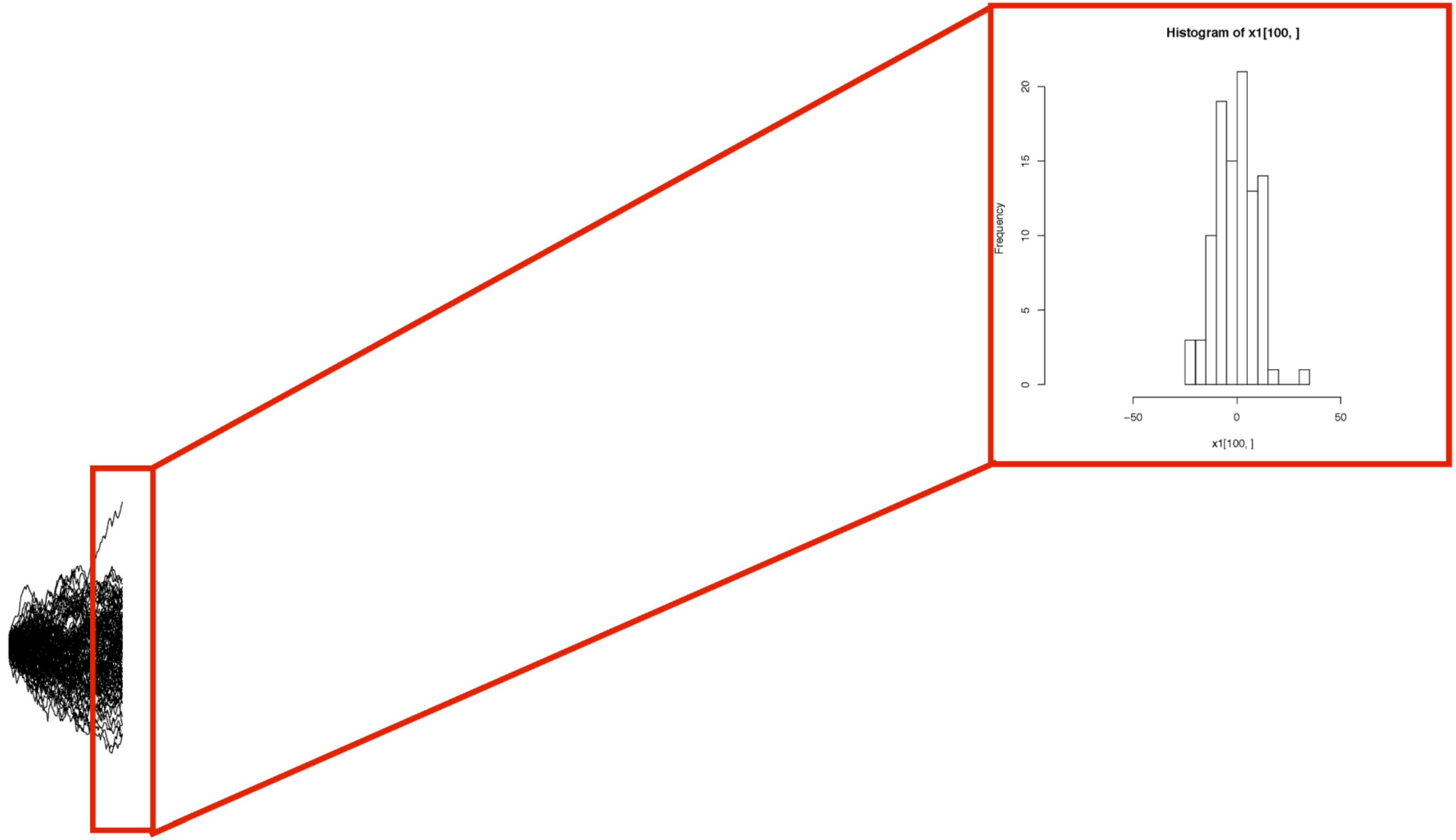
Three facts describe Brownian motion

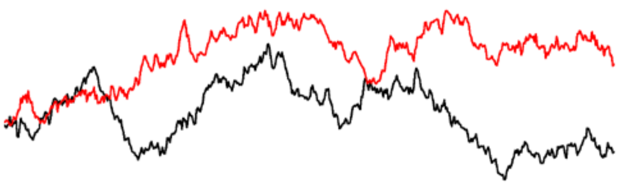
- Let $W(t)$ be the value of the character at time t.
  1. $E[W(t)] = W(0)$
  2. Successive changes are independent.
  3. $W(t) \sim N(W(0), \sigma^2 t)$

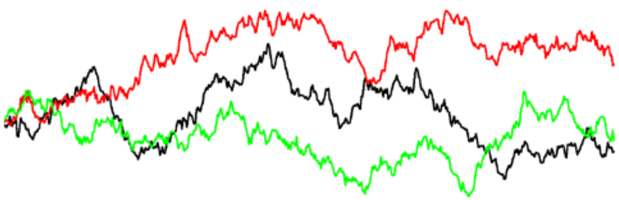
Parameters of the Brownian model

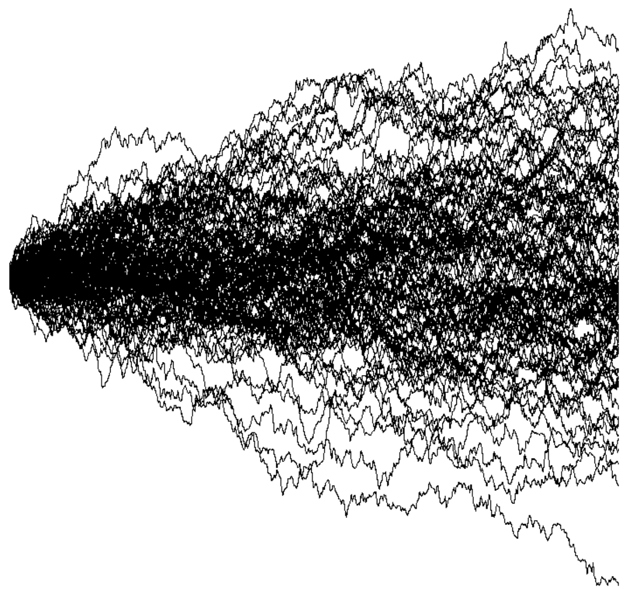
- The Brownian motion model has only two parameters:
  1. Θ, the starting value; W(0) = Θ.
  2. σ^2, the rate of accumulation of variance through time.

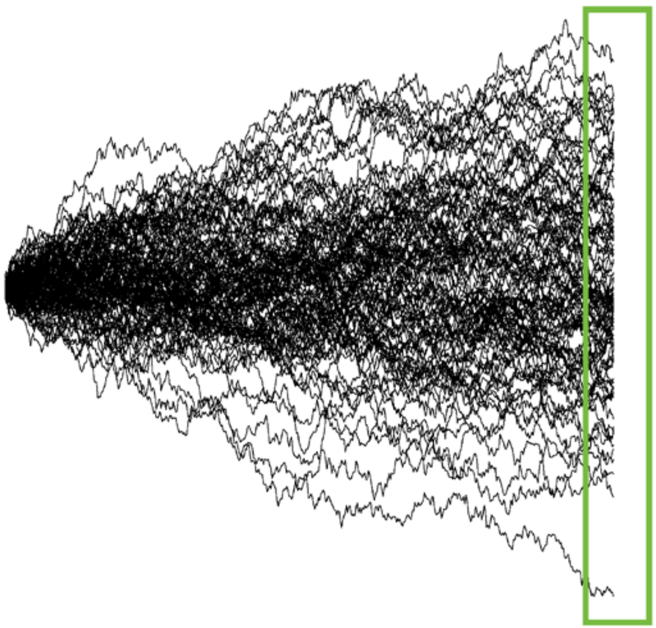


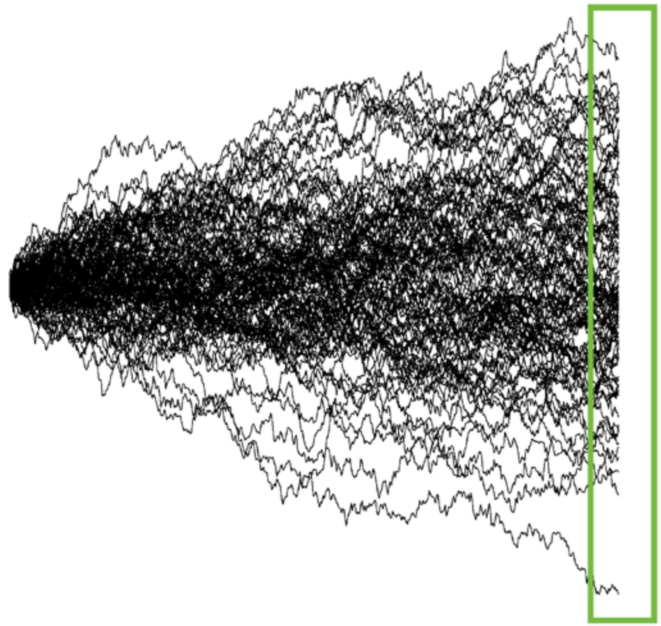

Histogram of x1[100,]
Frequency
0
5
10
15
20
−50
0
50
x1[100,]

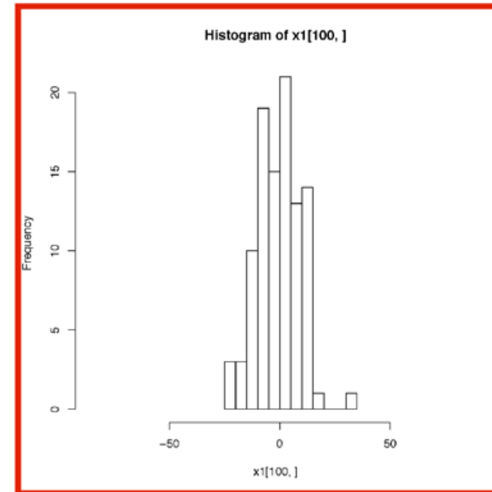


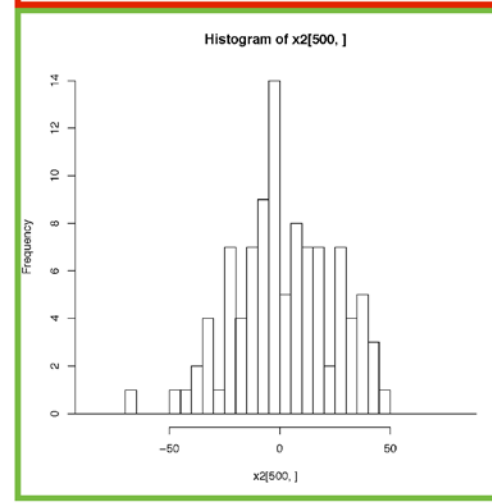


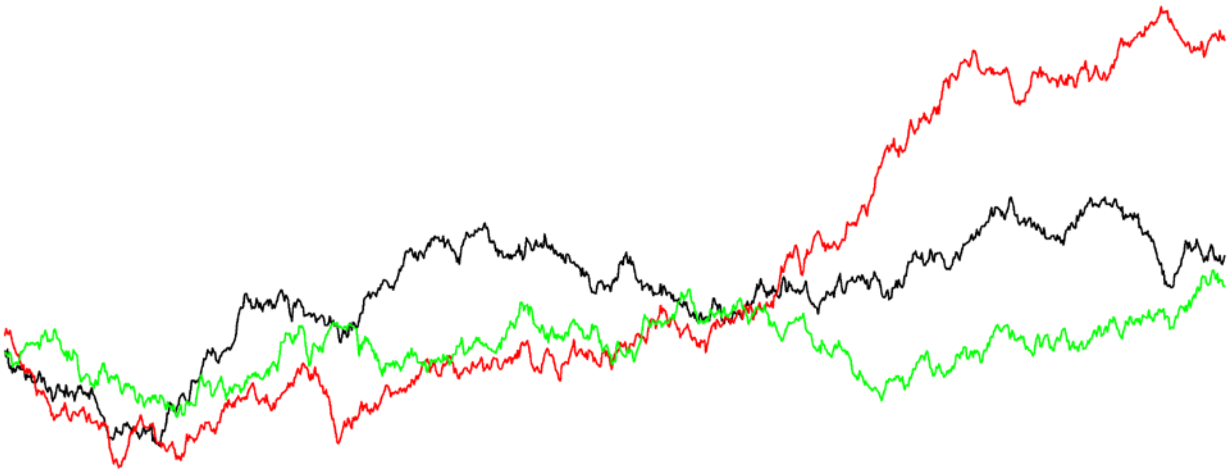


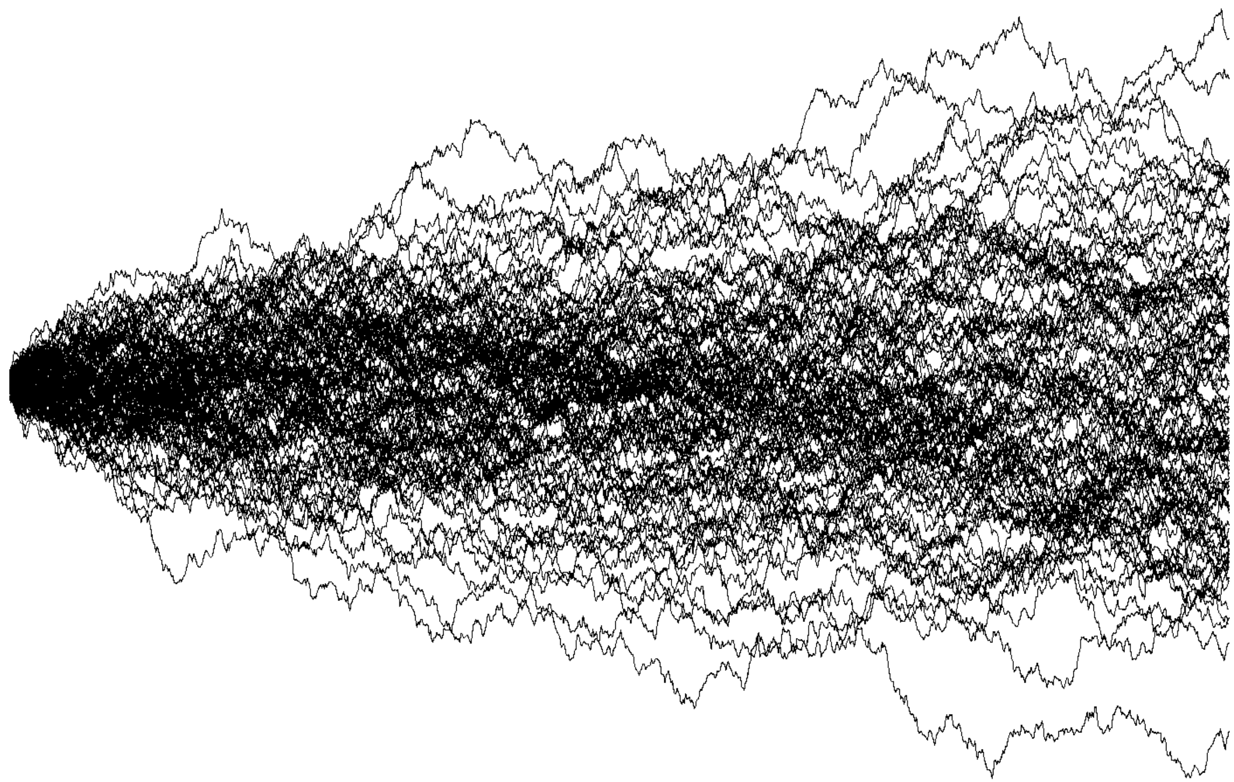


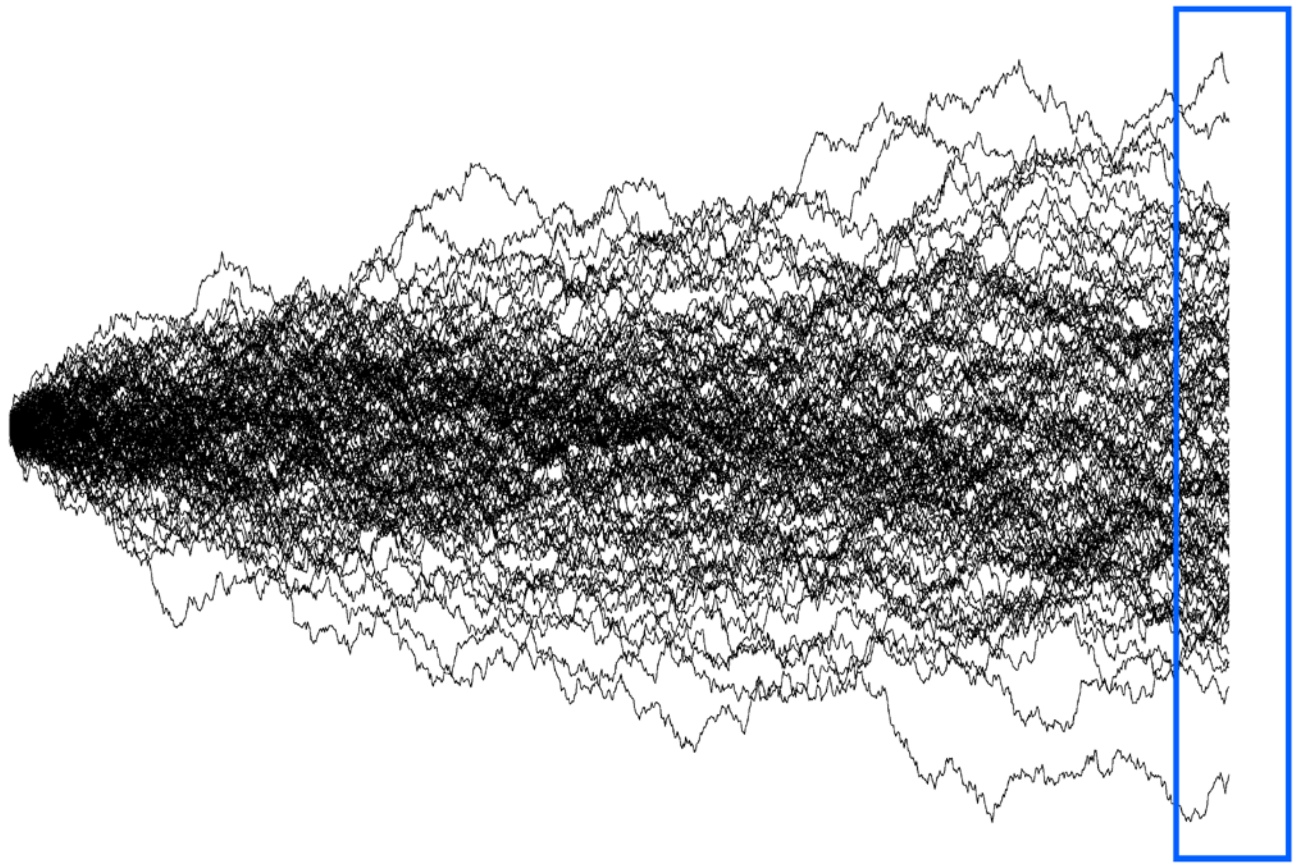


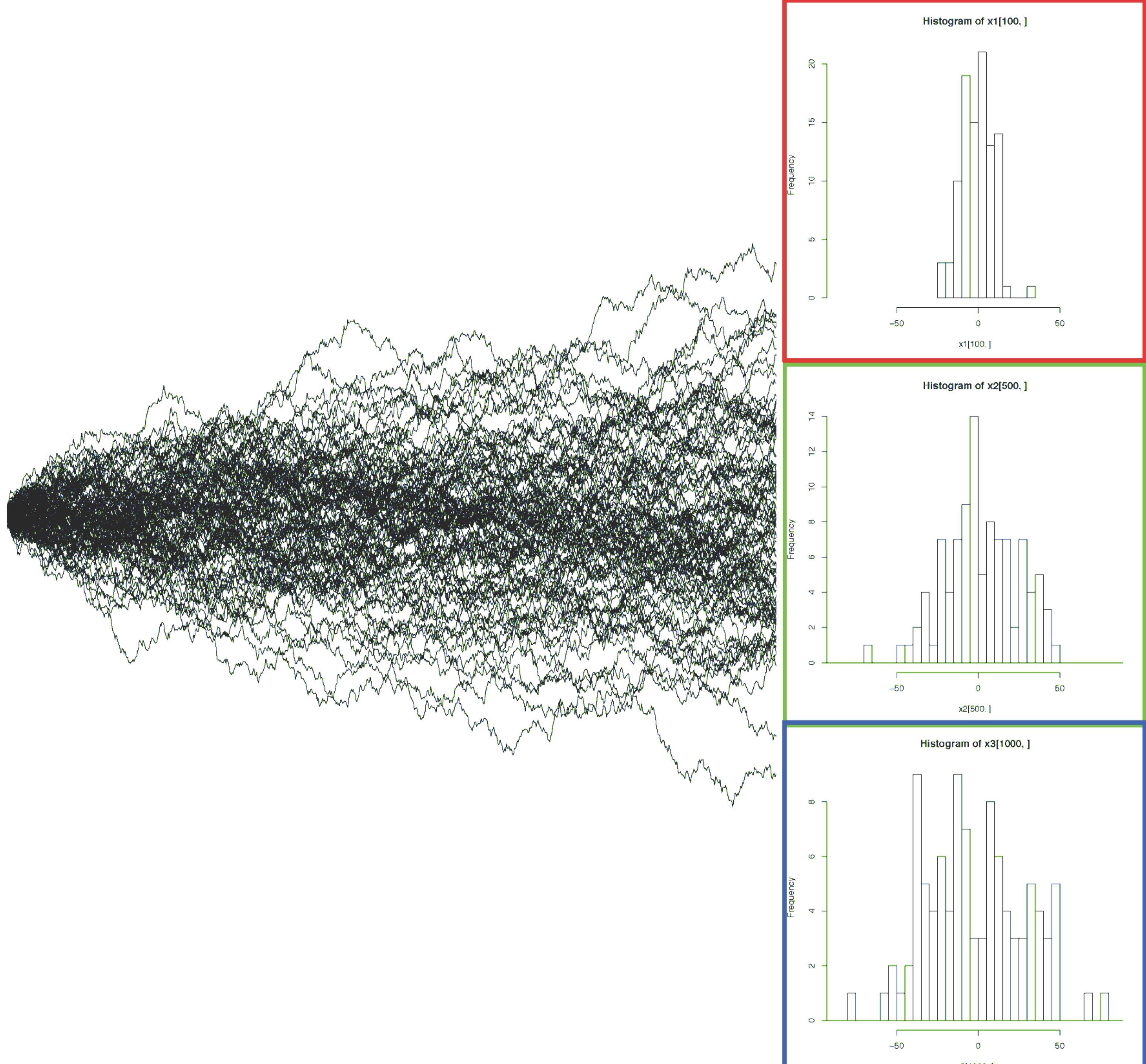

Histogram of x1[100,]
Frequency
0
5
10
15
20
−50
0
50
x1[100,]

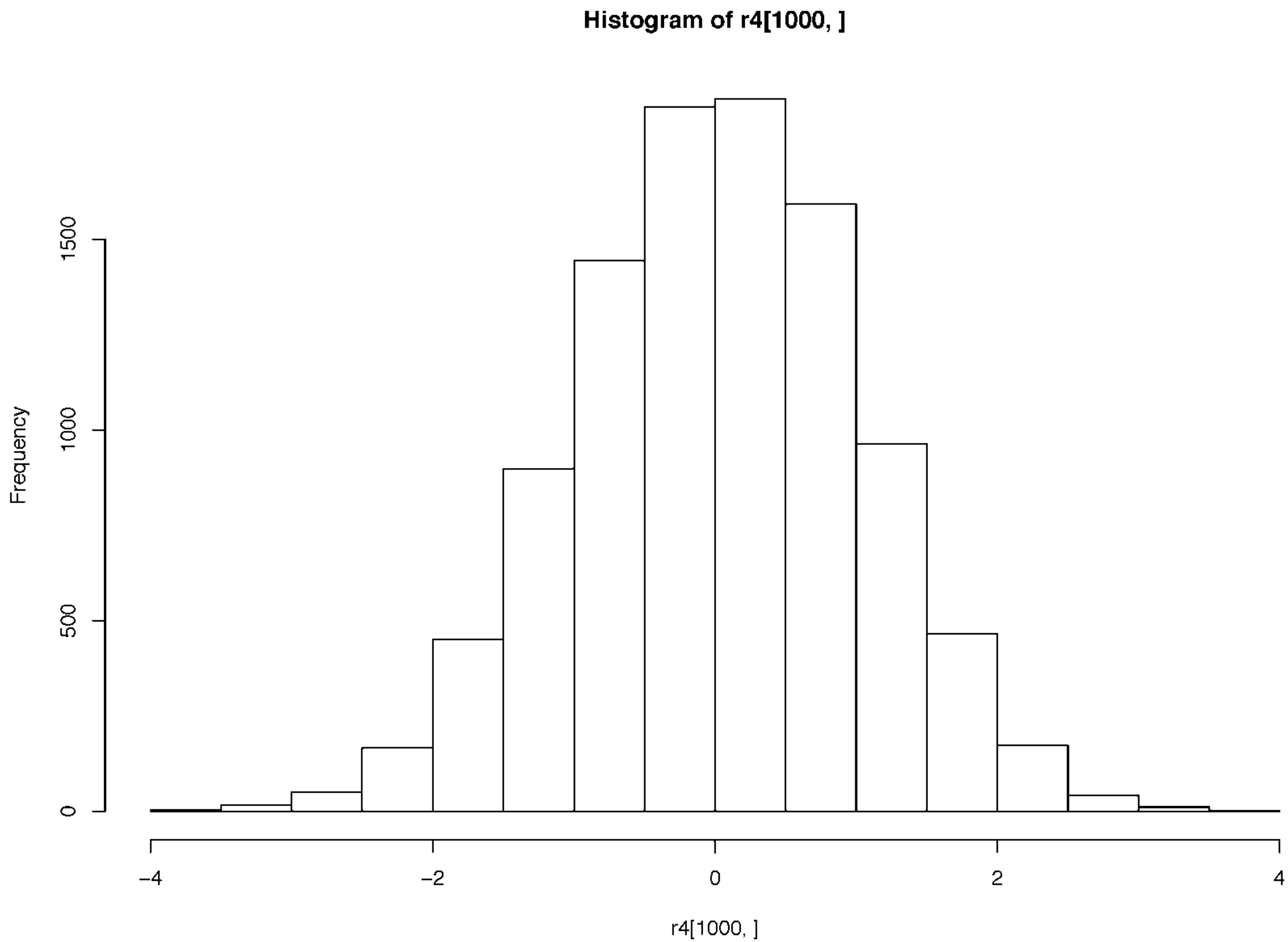

Histogram of x2[500,]
Frequency
0
2
4
6
8
10
12
14
−50
0
50
x2[500,]



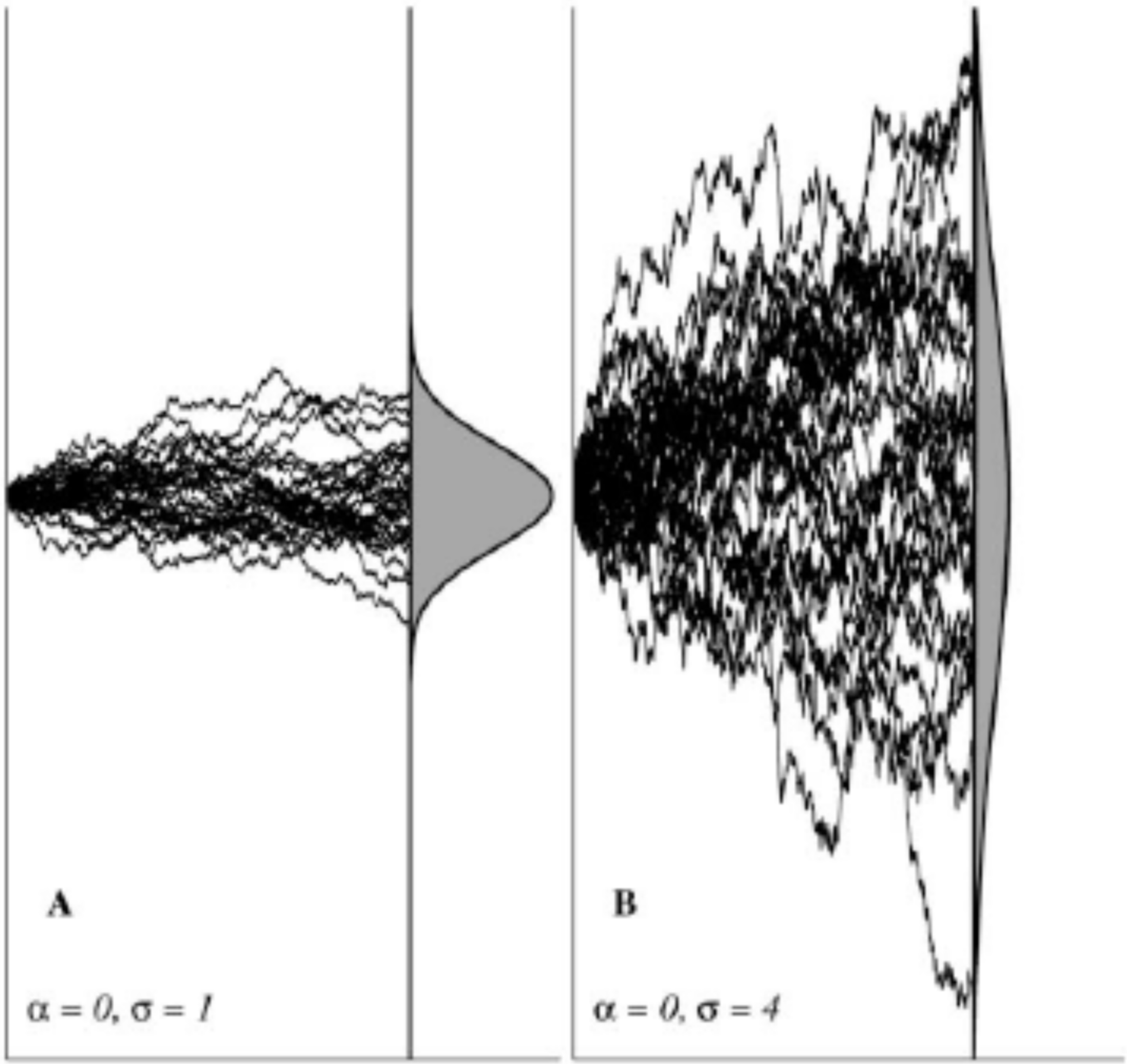





Histogram of x1[100,]
Frequency
x1[100,]
Histogram of x2[500,]
Frequency
x2[500,]
Histogram of x3[1000,]
Frequency



Histogram of r4[1000,]
Frequency
0
500
1000
1500
−4
−2
0
2
4
r4[1000,]



A
α = 0, σ = 1
B
α = 0, σ = 4

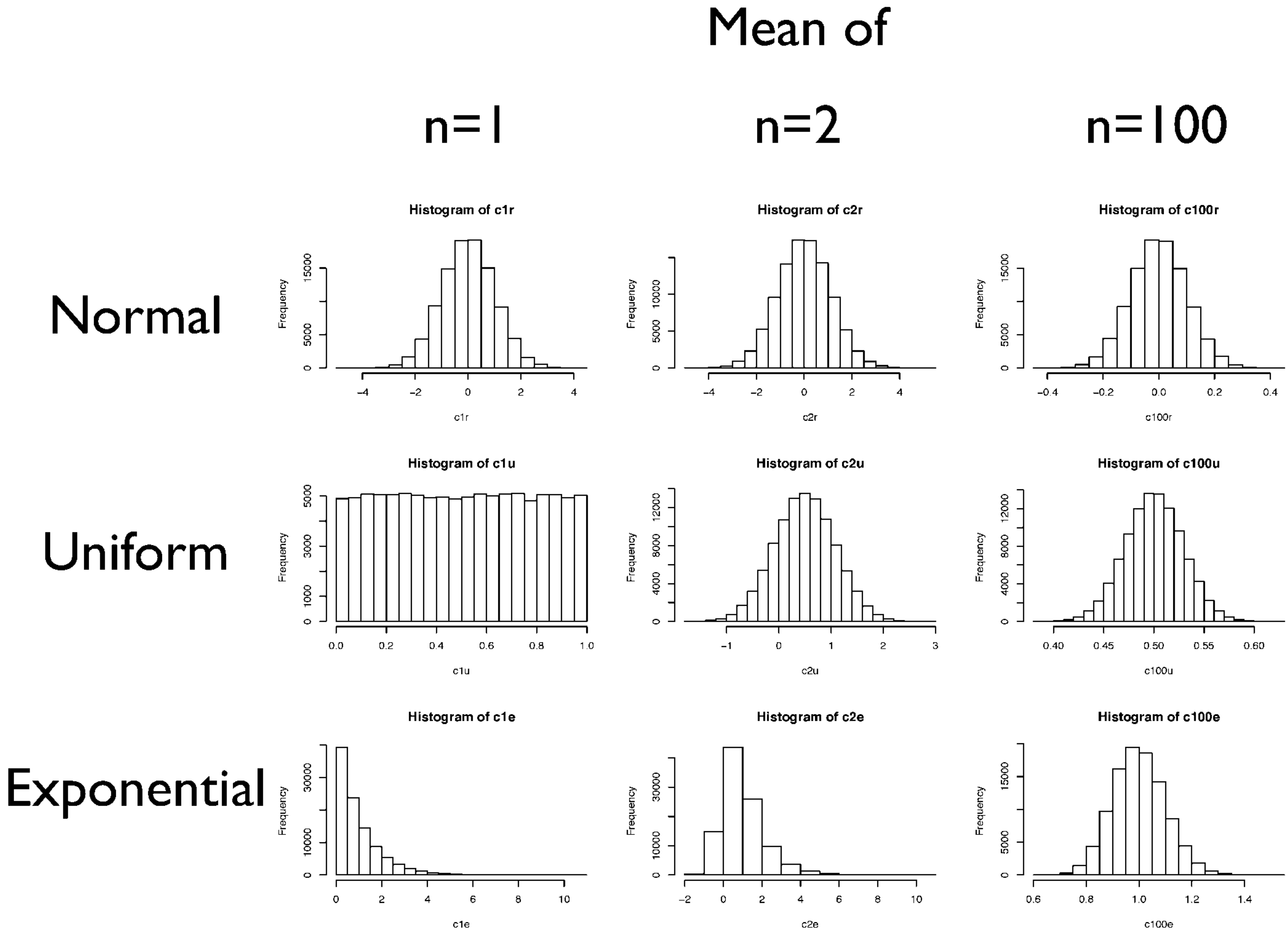
Outline

1. What is Brownian motion?
2. When might characters evolve by Brownian motion.
3. How can we simulate Brownian motion evolution on trees?

A physical model for BM

Why Normal?

- BM can be used to describe motion that results from the combination of a large number of independent weak forces
- Adding many small independent variables result in normal distributions, no matter the original distribution (Central limit theorem)


Mean of
n=1
n=2
n=100
Normal
Uniform
Exponential
Histogram of c1r
Histogram of c2r
Histogram of c100r
Histogram of c1u
Histogram of c2u
Histogram of c100u
Histogram of c1e
Histogram of c2e
Histogram of c100e
Frequency
c1r
c2r
c100r
c1u
c2u
c100u
c1e
c2e
c100e

Evolution might approximate BM...

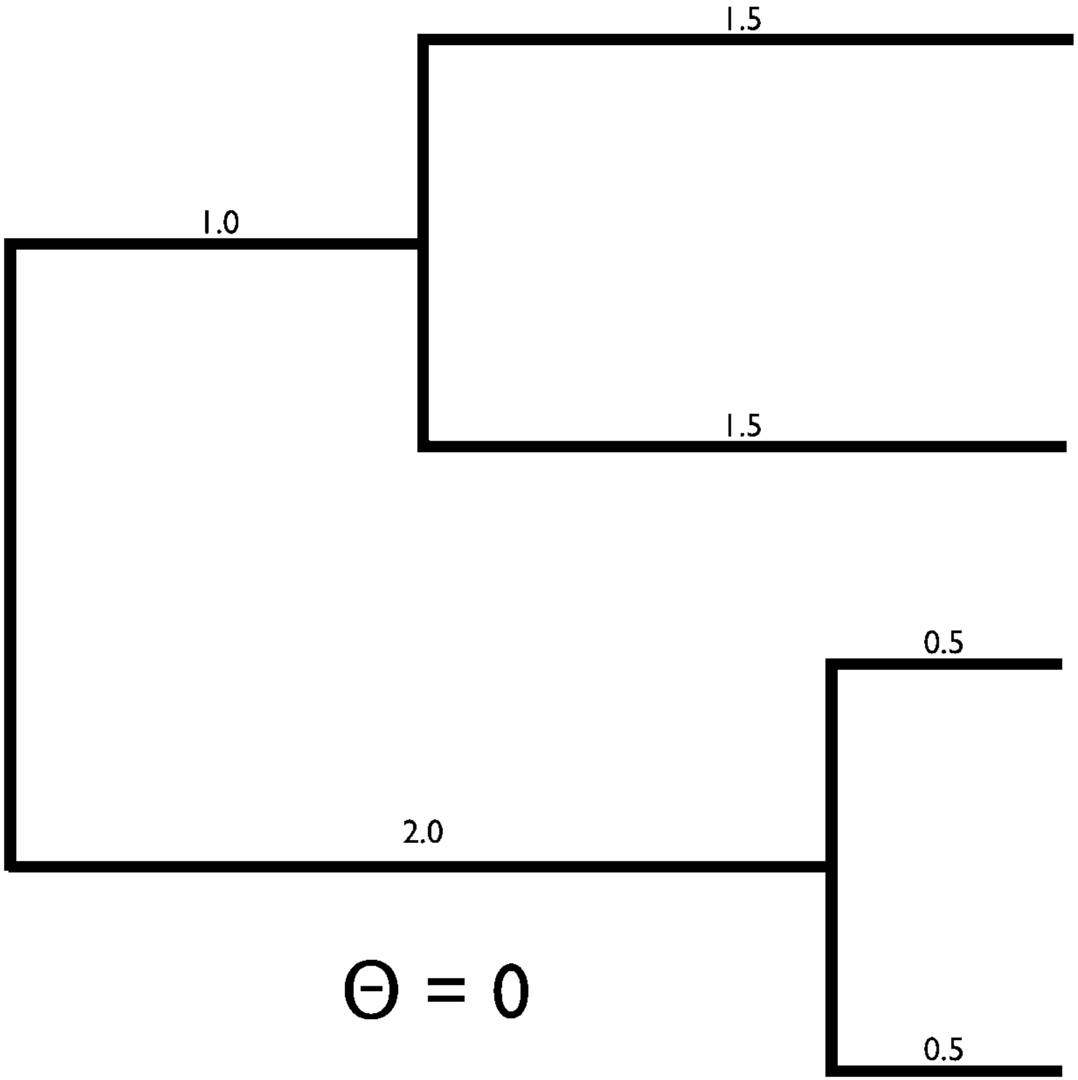
- Genetic drift
- Random punctuated change
- Selection that is weak relative to the time interval considered
- Selection that changes randomly through time

Outline - BM

- What is Brownian motion?
- When might characters evolve in a Brownian-like way?
- Simulating Brownian motion on trees

Simulating BM

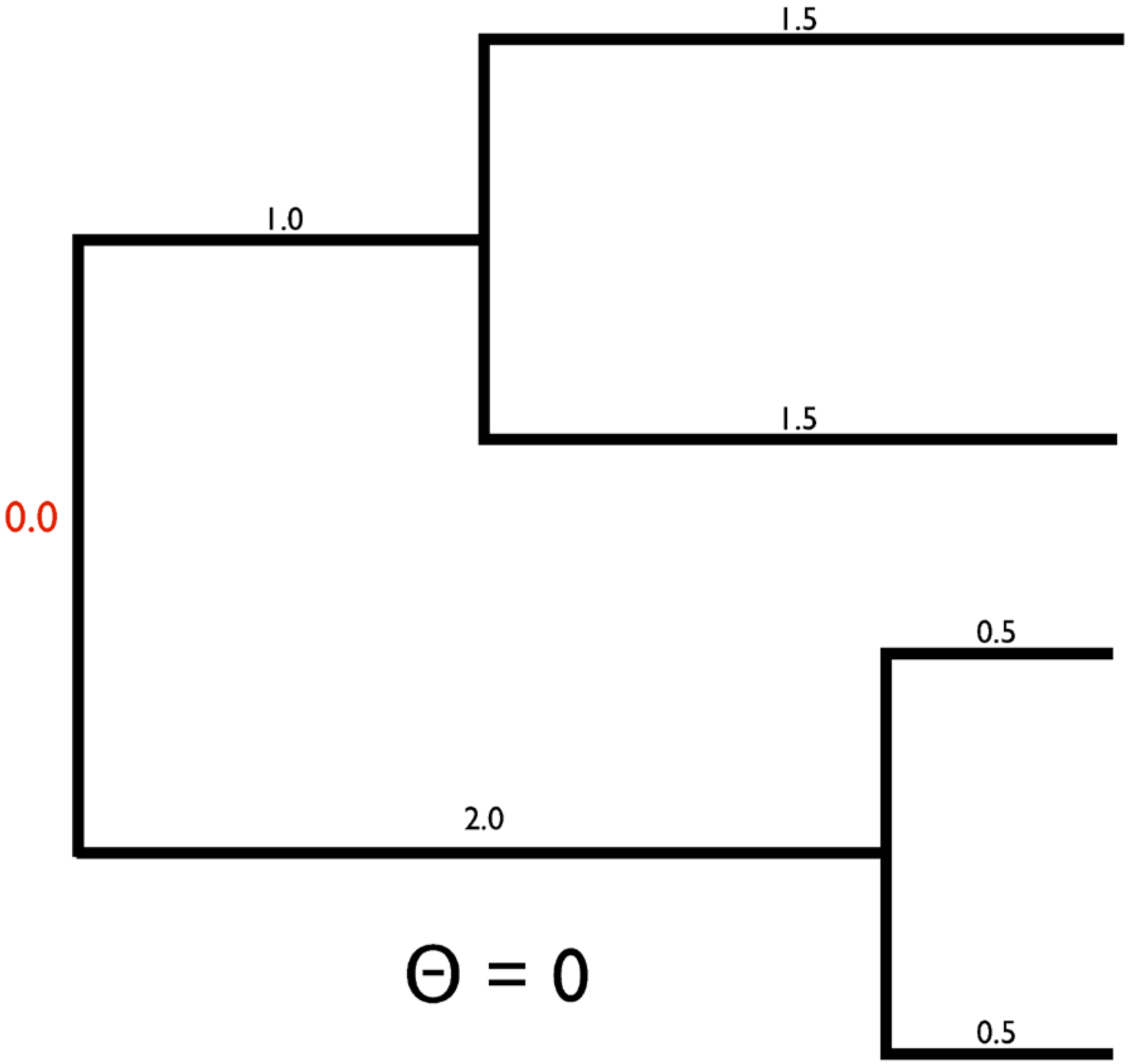
- Simulating Brownian motion involves drawing values from normal distributions
- Variance of the distribution depends on σ^2 and t
- Values along adjacent branches are added from the root to the tips of the tree



$\Theta = 0$

$\sigma^2 = 1.0$

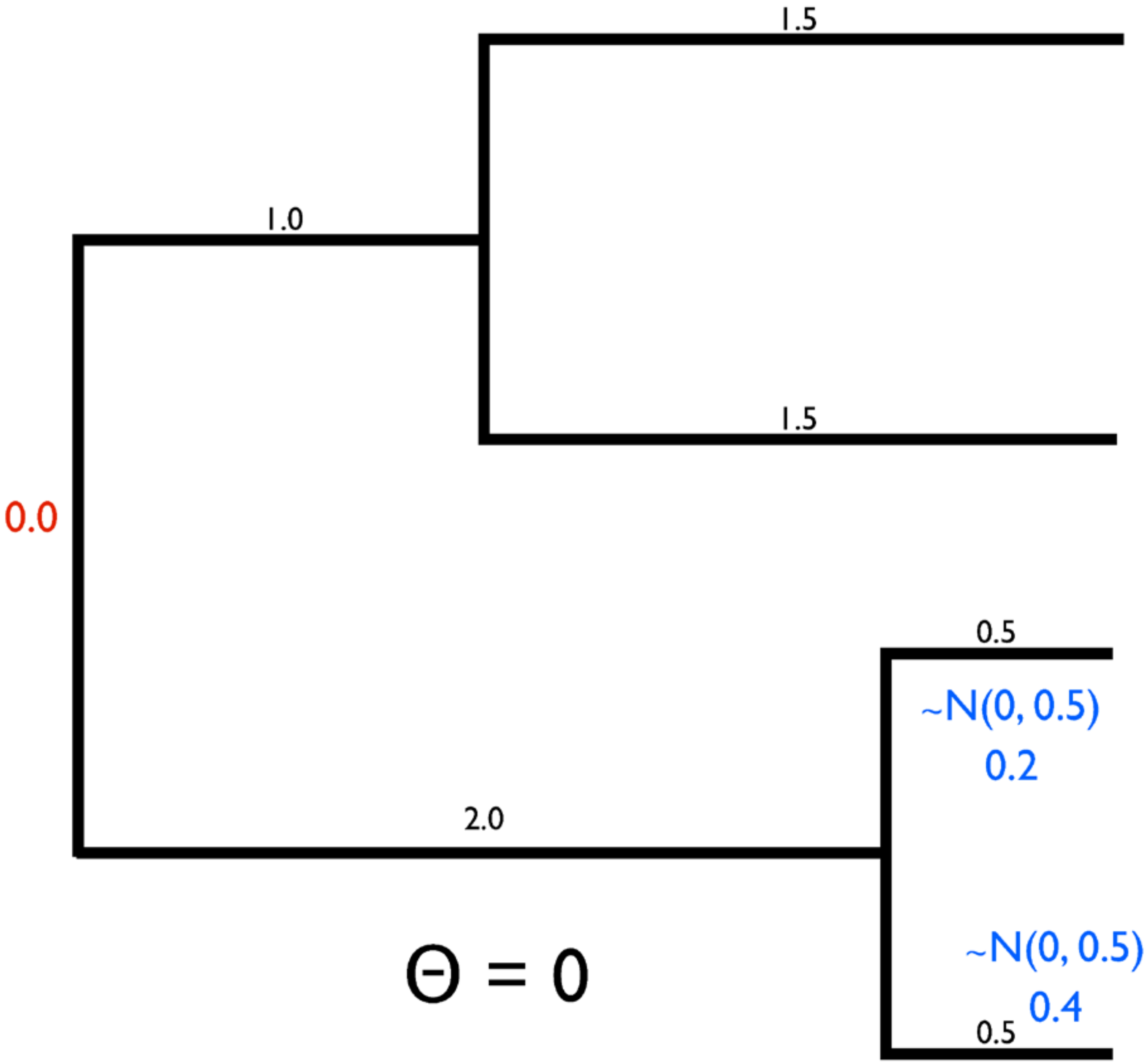
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



$\Theta = 0$

$\sigma^2 = 1.0$

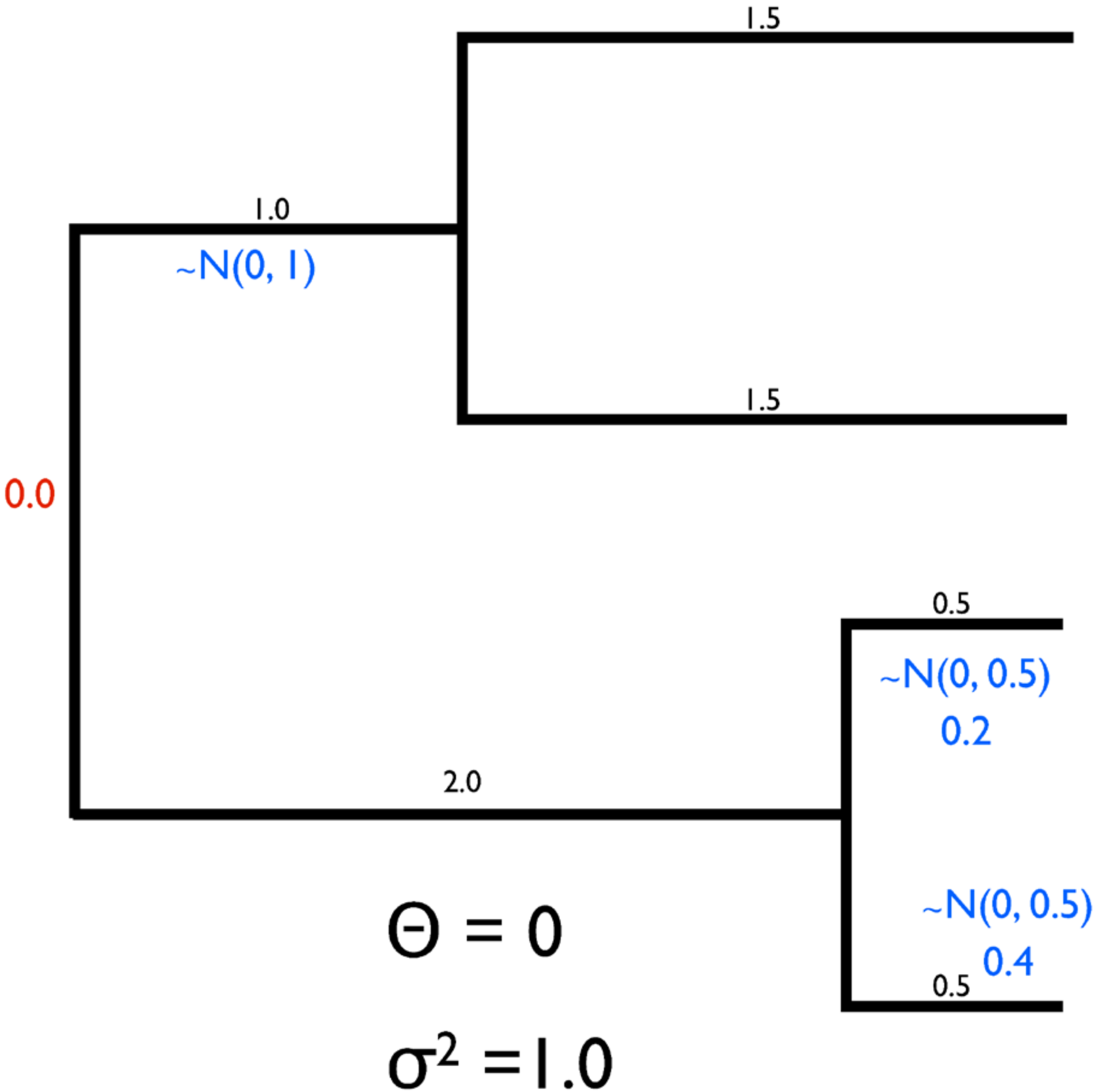
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



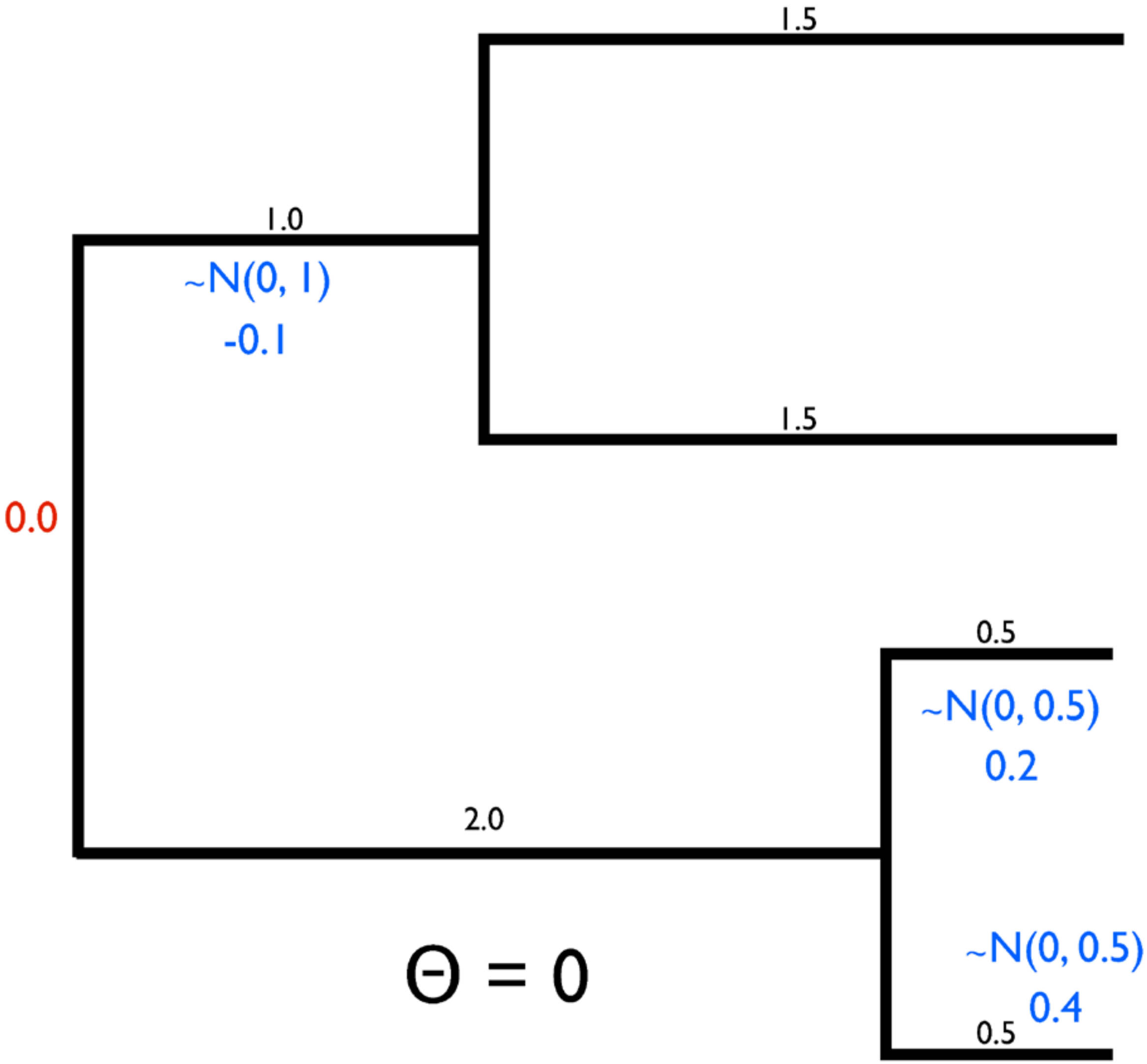
$\Theta = 0$

$\sigma^2 = 1.0$

1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



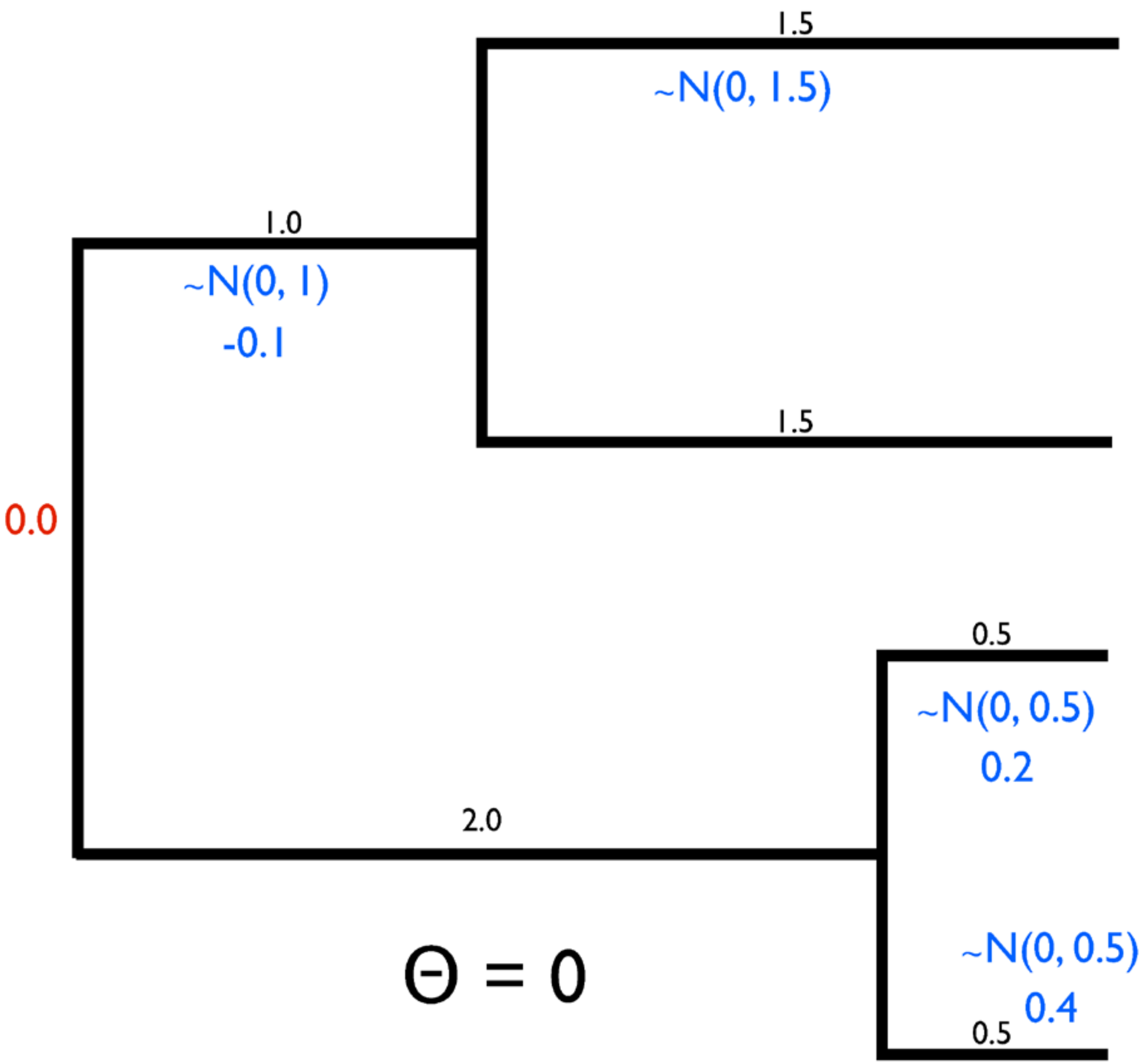
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

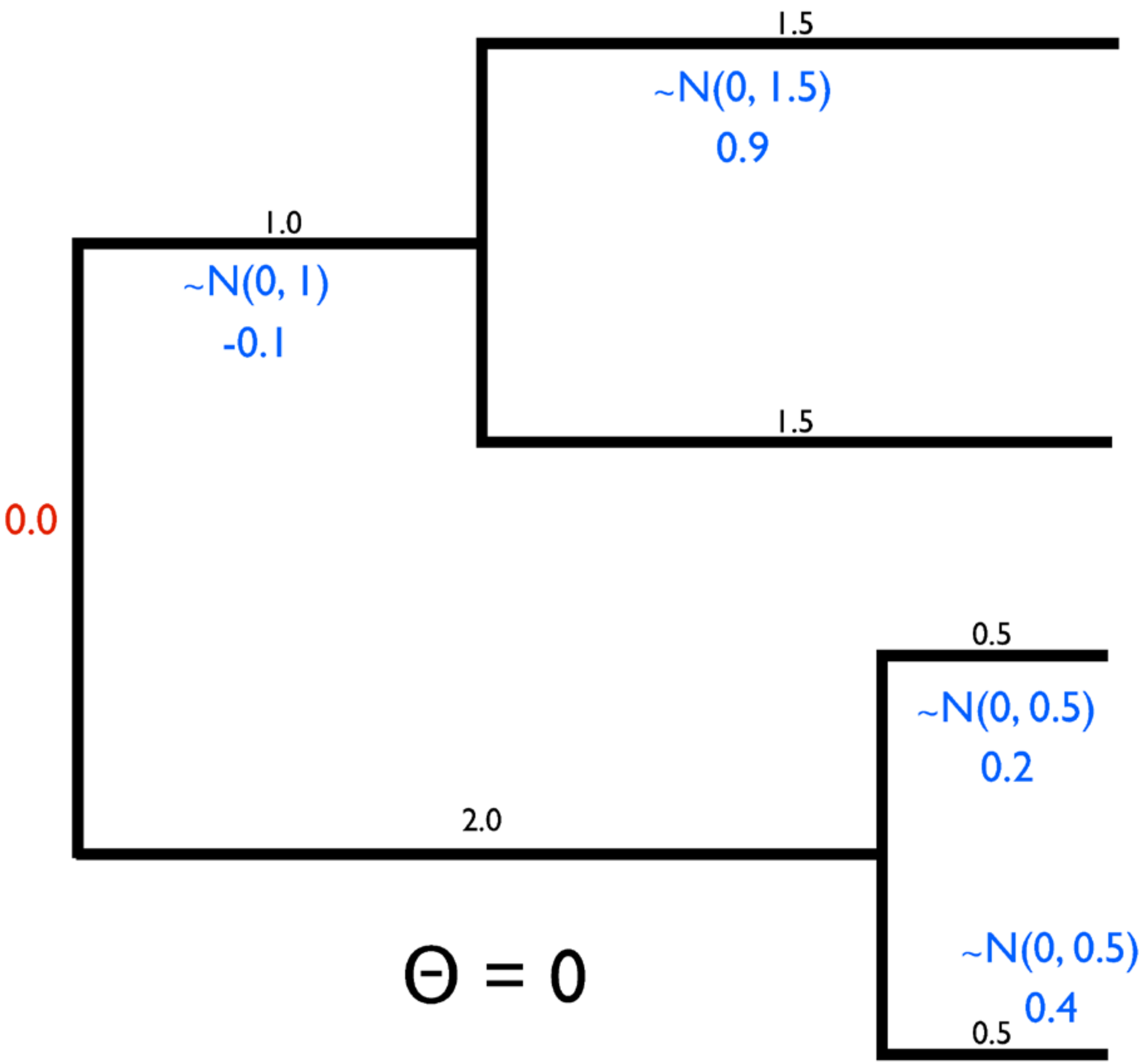
$\sigma^2 = 1.0$



$\Theta = 0$

$\sigma^2 = 1.0$

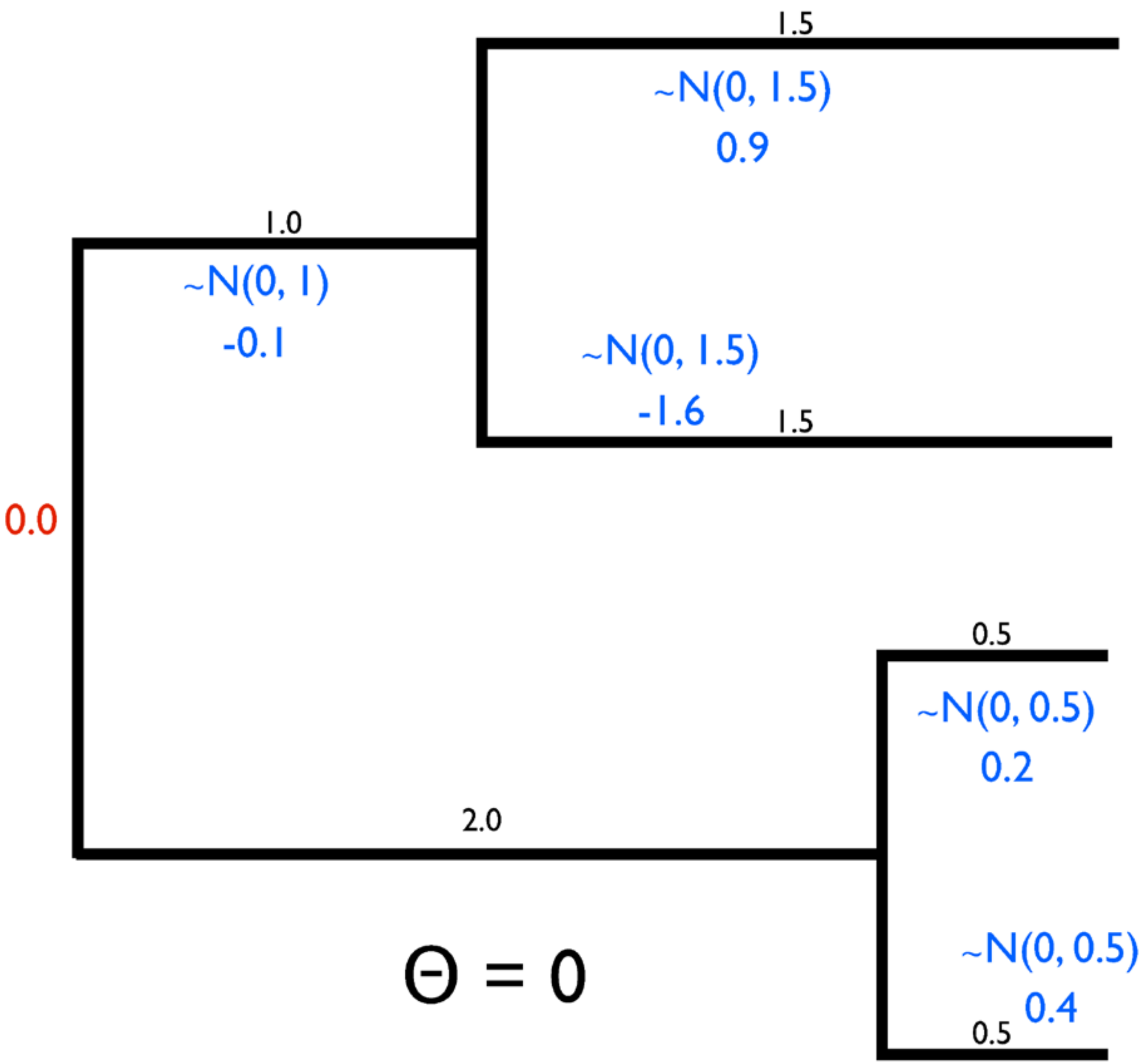
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

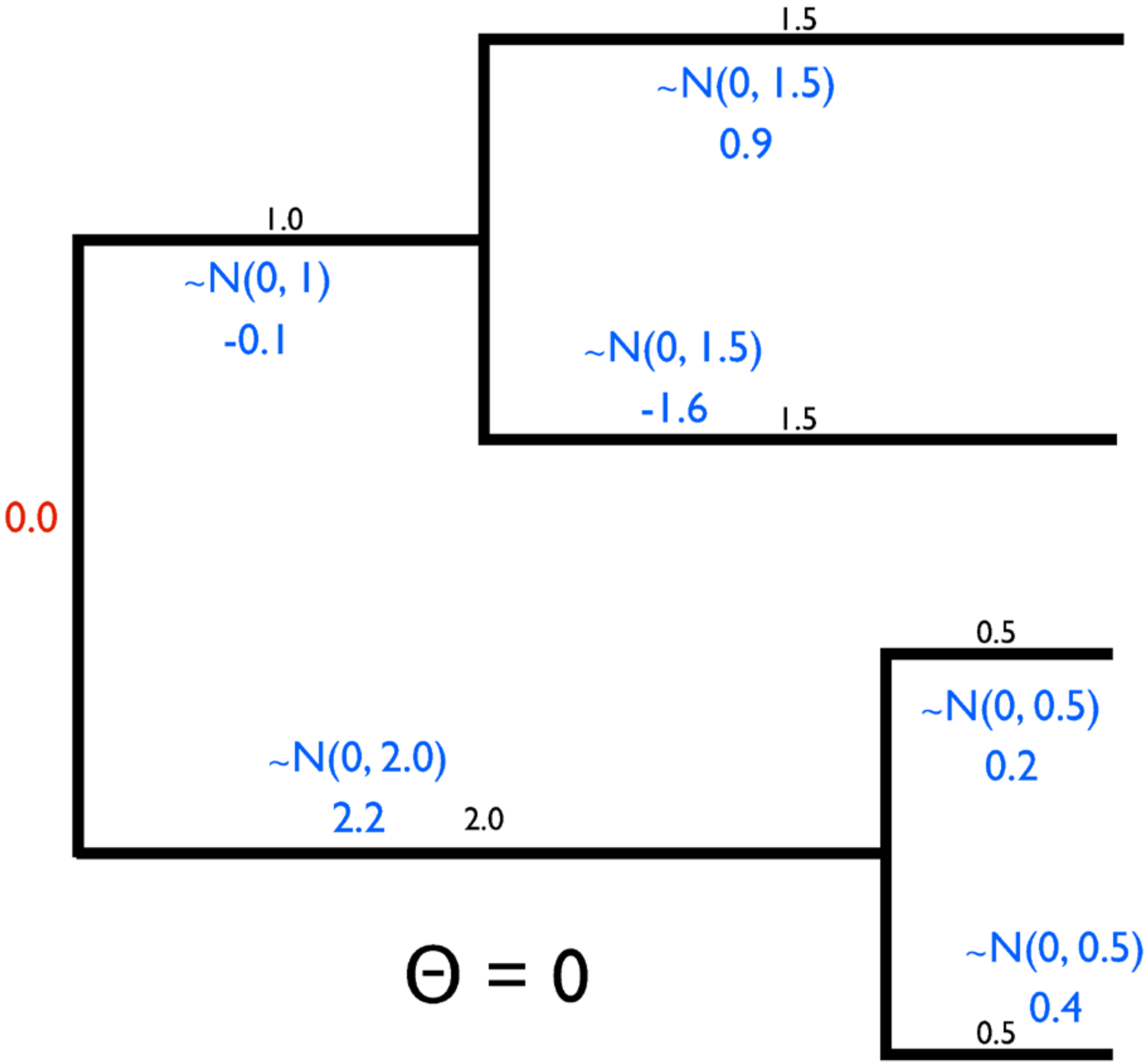
$\sigma^2 = 1.0$



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

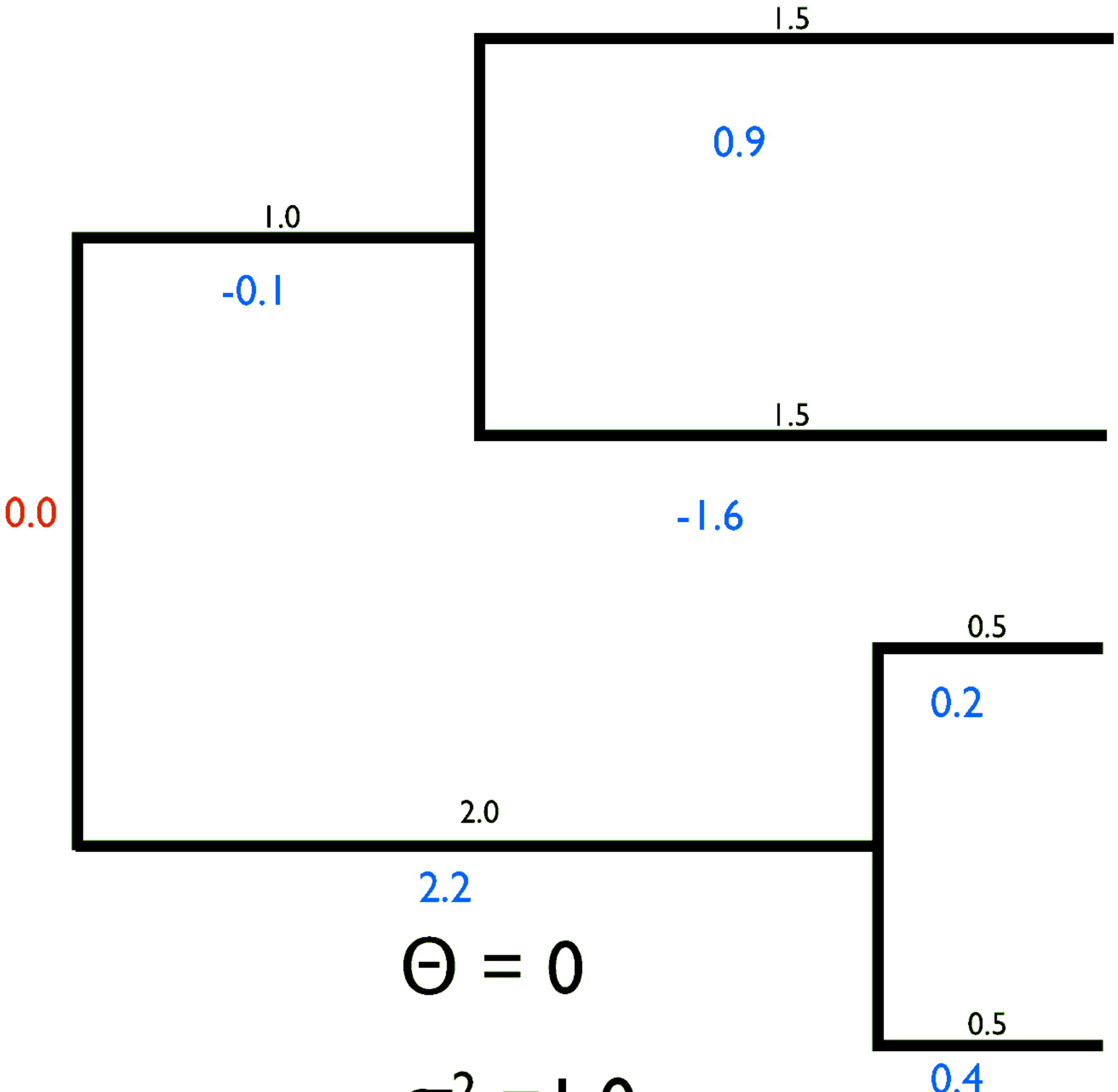
$\sigma^2 = 1.0$



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

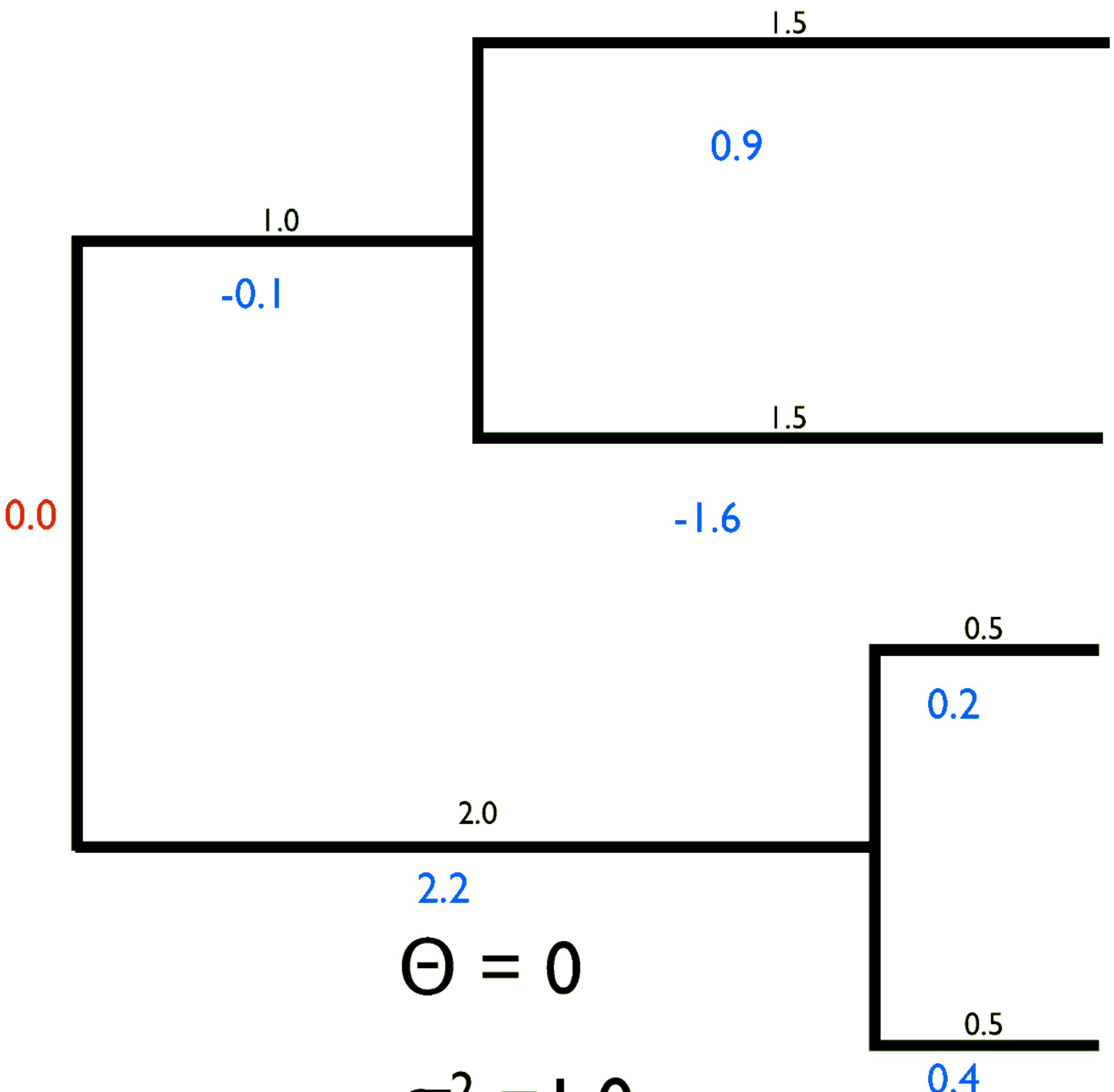
$\sigma^2 = 1.0$



$\Theta = 0$

$\sigma^2 = 1.0$

1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

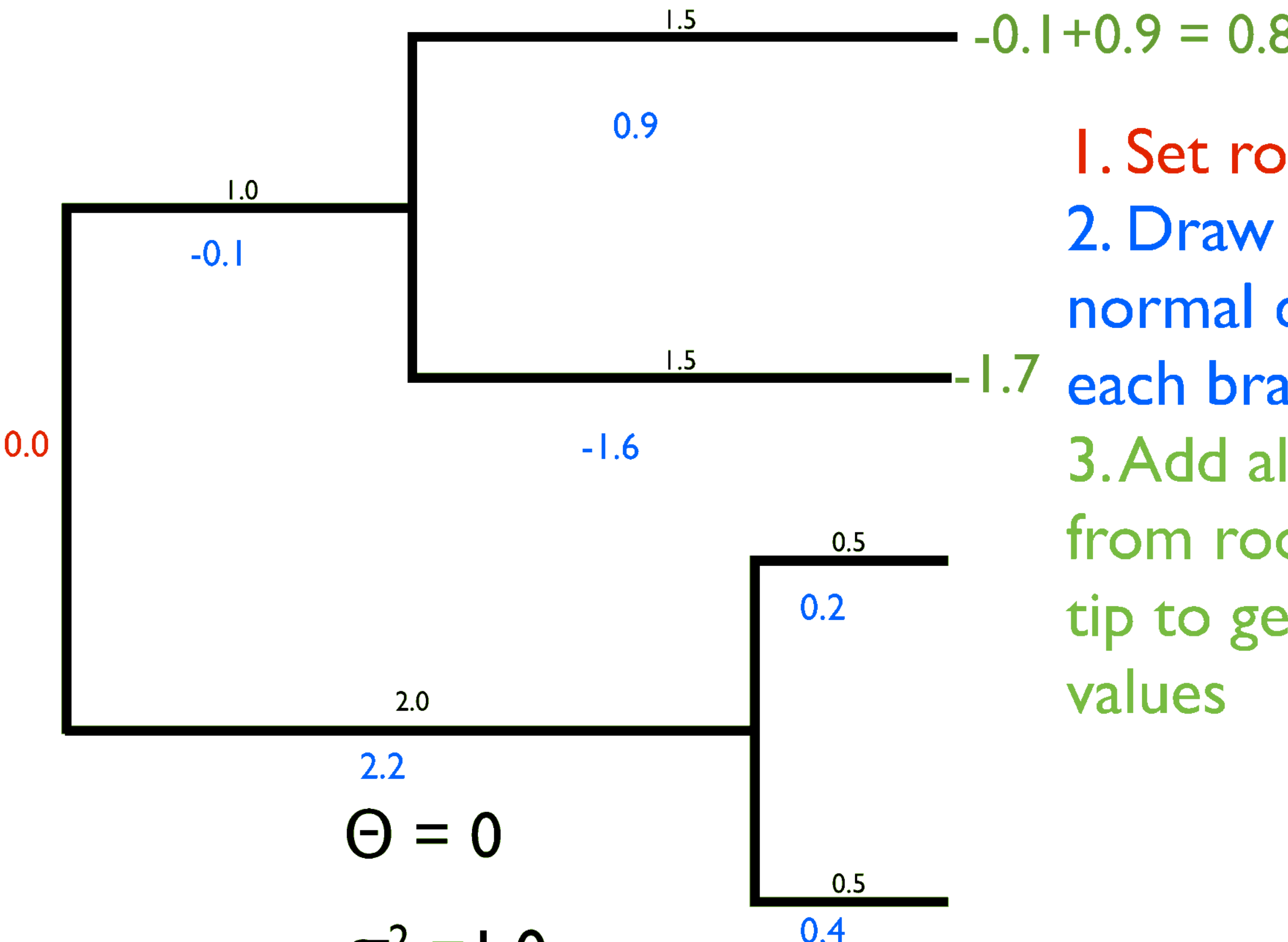


$-0.1+0.9 = 0.8$

1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

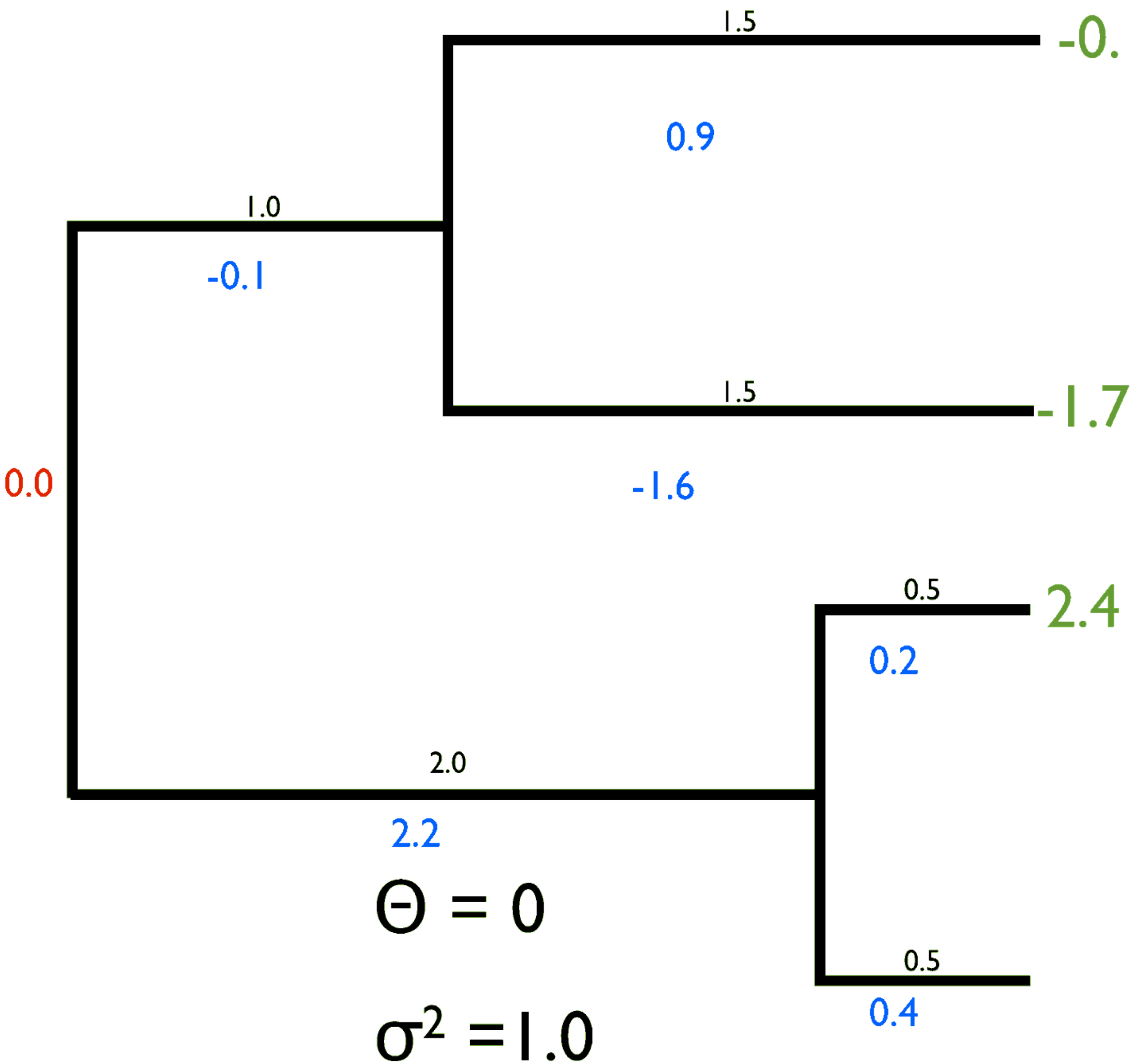
$\sigma^2 = 1.0$



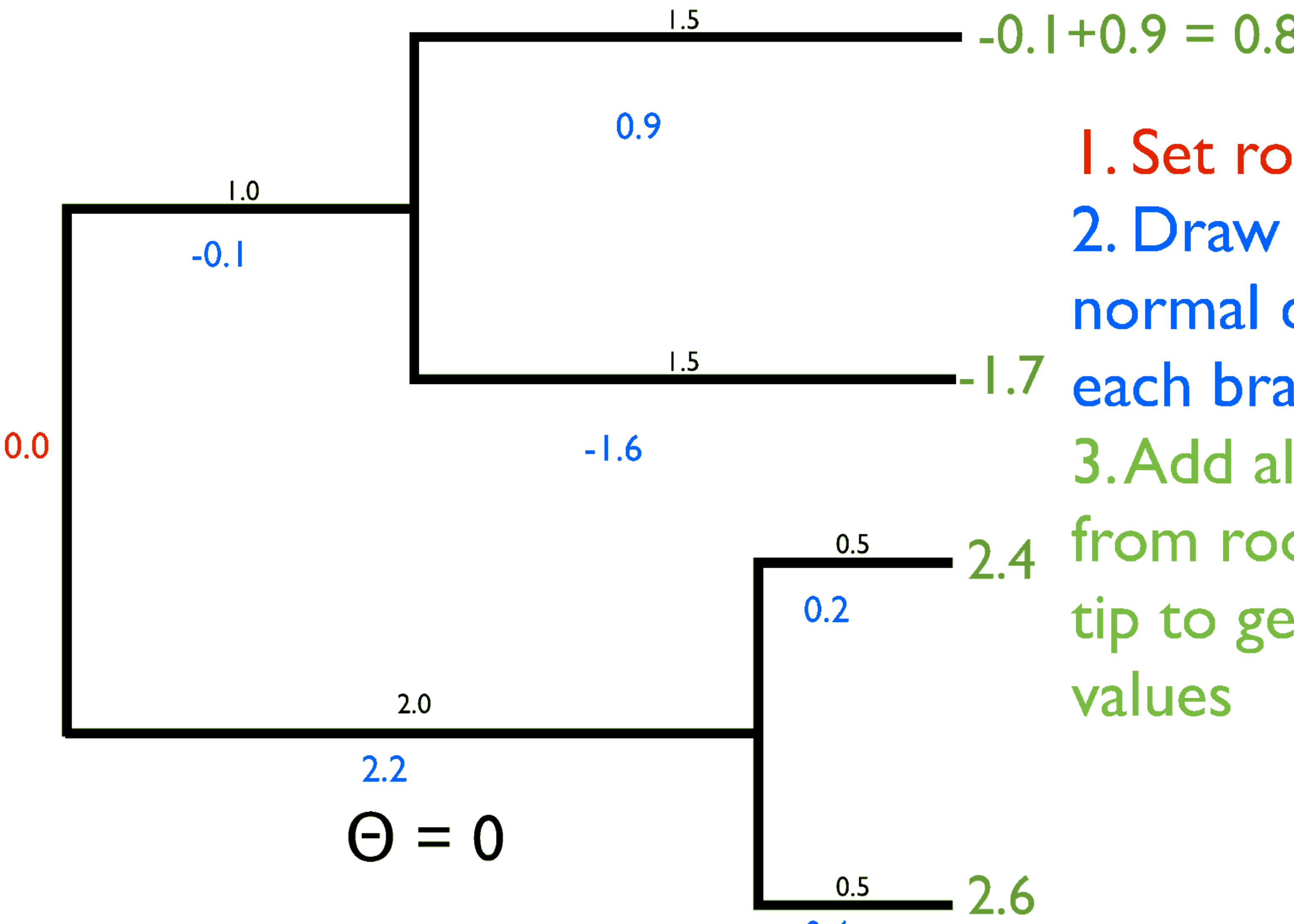
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

$\sigma^2 = 1.0$



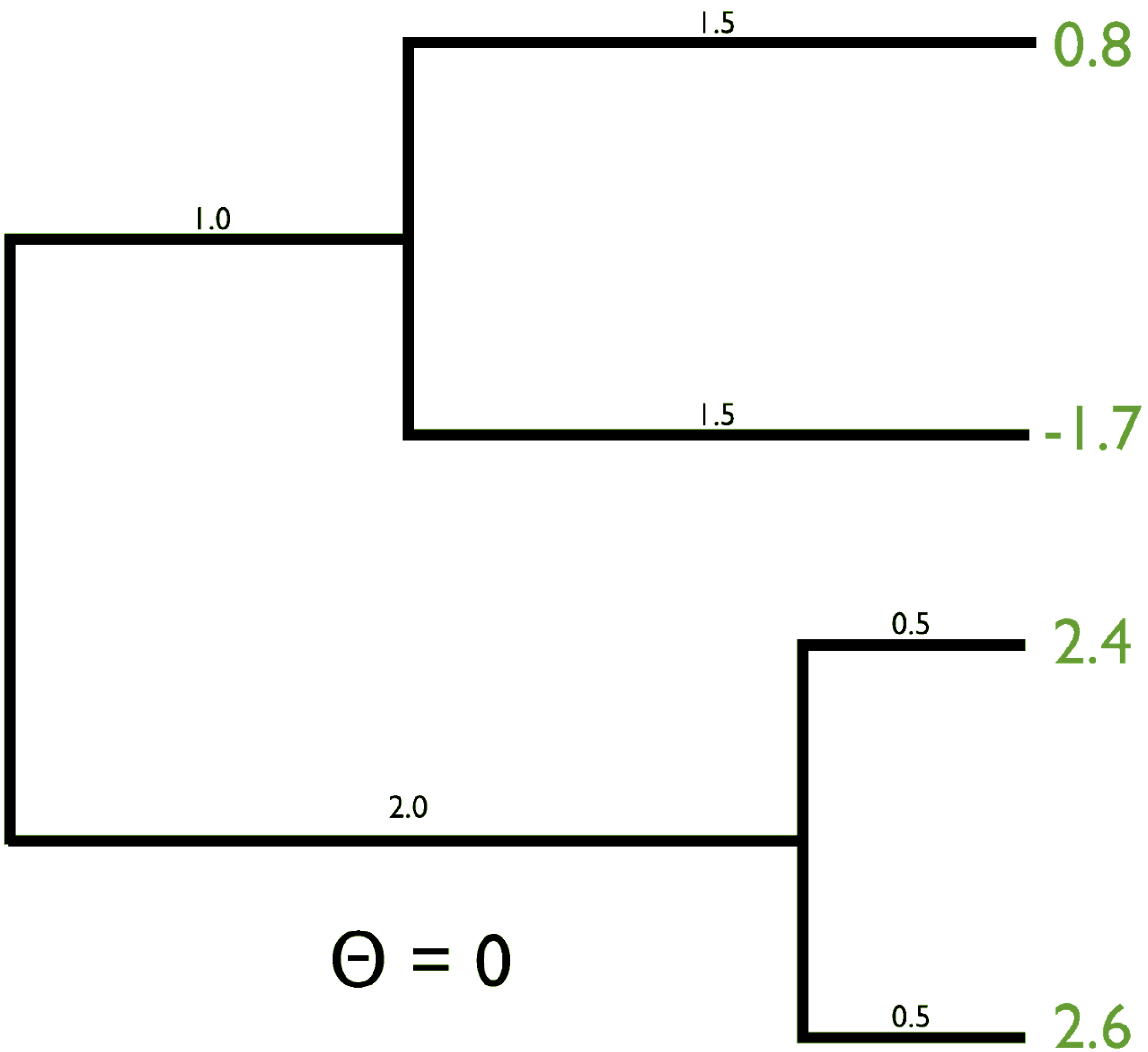
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

$\Theta = 0$

$\sigma^2 = 1.0$



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

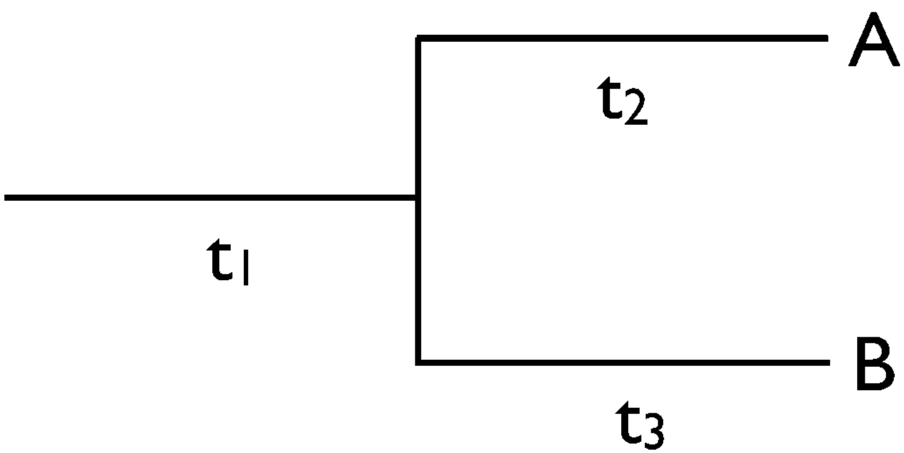
$\Theta = 0$

$\sigma^2 = 1.0$

Properties of BM on trees

- Variance increases with both σ^2 and t
- Expected (mean) value of any tip is always Θ
- Closely related species tend to be similar (they covary)

How do they covary?



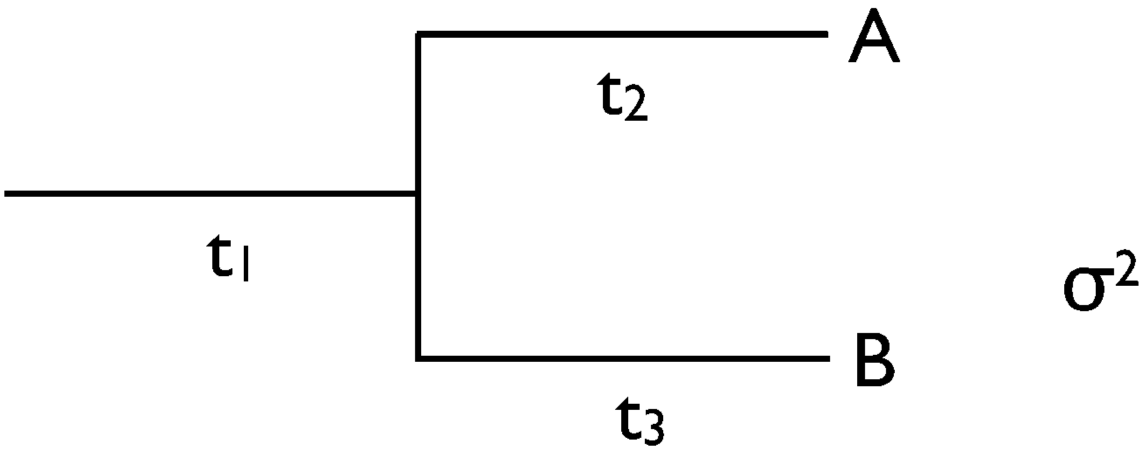
σ^2

var(A)

cov(A,B)

var(B)

How do they covary?

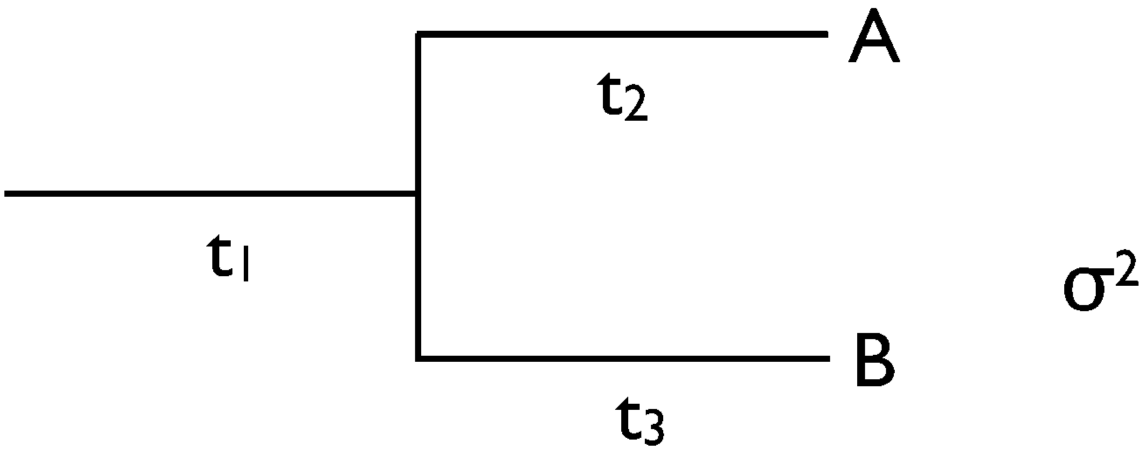


var(A) $=\sigma^2(t_1+t_2)$

cov(A,B)

var(B)

How do they covary?

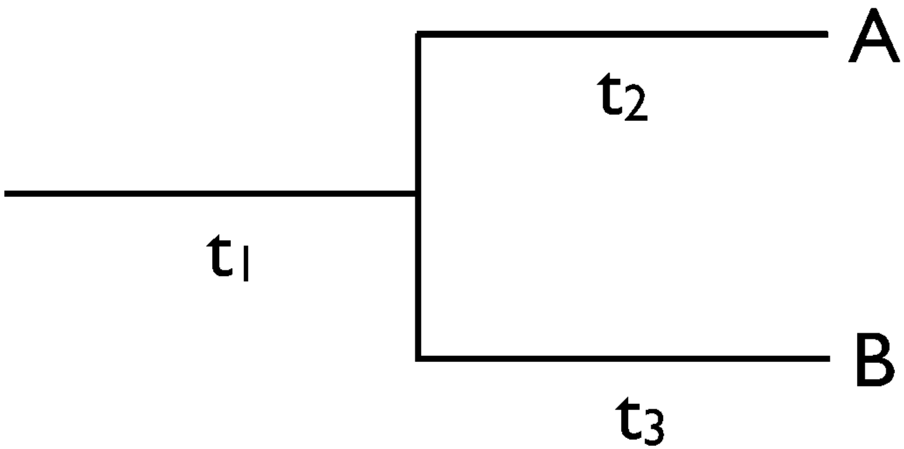


$var(A) = \sigma^2(t_1+t_2)$

cov(A,B)

$var(B) = \sigma^2(t_1+t_3)$

How do they covary?



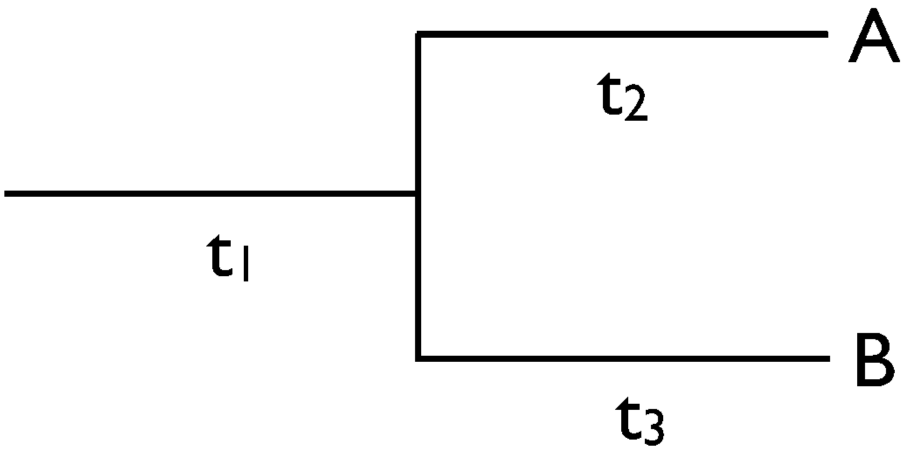
σ^2

$\mathrm{var}(A) = \sigma^2(t_1+t_2)$

$\mathrm{cov}(A,B) = \sigma^2(t_1)$

$\mathrm{var}(B) = \sigma^2(t_1+t_3)$

How do they covary?



variance-covariance matrix

$$\sigma^2 \begin{bmatrix} t_1+t_2 & t_1 \\ t_1 & t_1+t_3 \end{bmatrix}$$

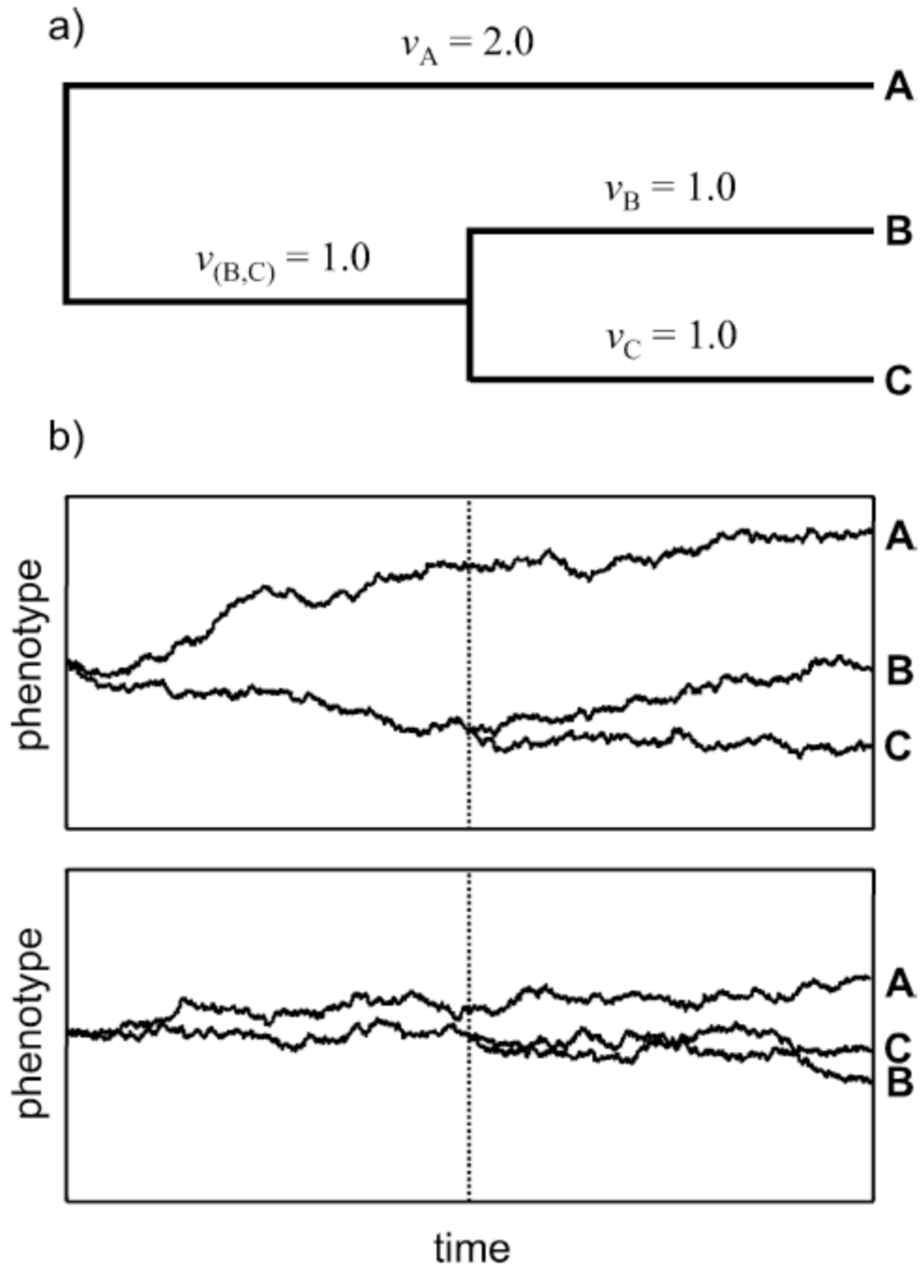
$\text{var}(A) = \sigma^2(t_1+t_2)$

$\text{cov}(A,B) = \sigma^2(t_1)$

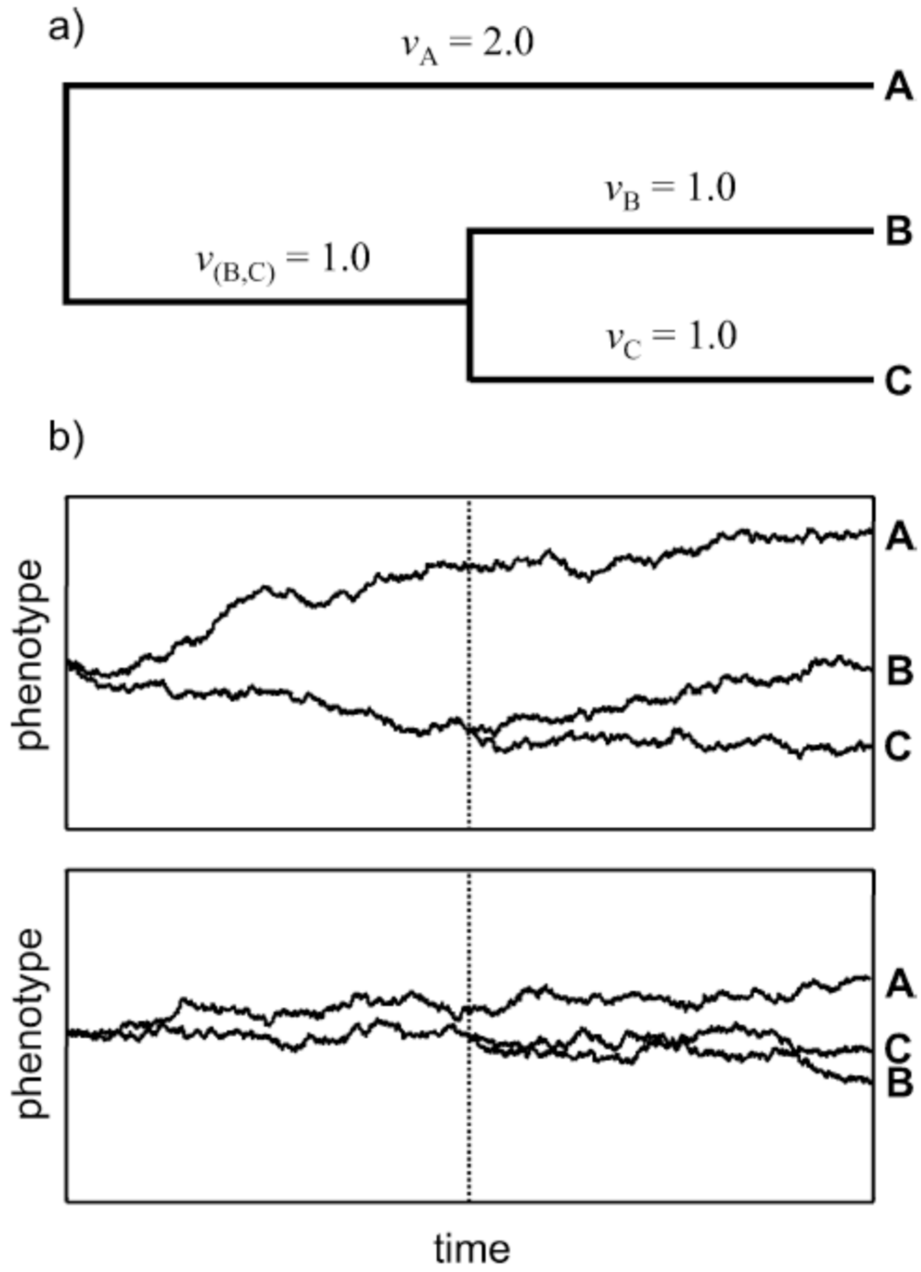
$\text{var}(B) = \sigma^2(t_1+t_3)$

General form

- Tip data follow a multivariate normal distribution with mean vector Θ and variance-covariance matrix where
- $var(i) = \sigma^2(d_i)$; d_i =distance from root to tip i
- $cov(i,j) = \sigma^2(c_{i,j})$; $c_{i,j}$ =shared path of tip i and j



from Revell et al. 2008



c)

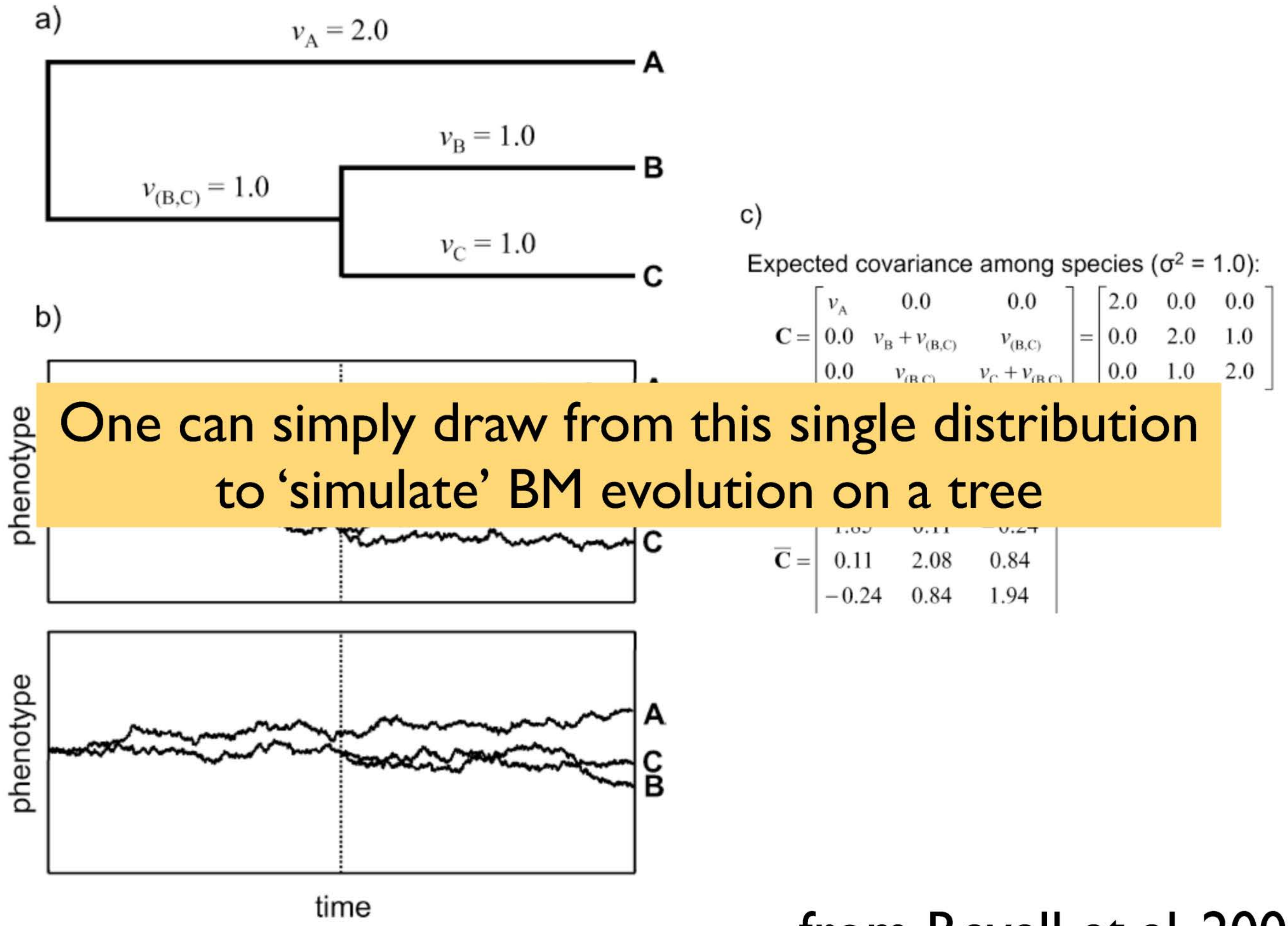
Expected covariance among species ($\sigma^2 = 1.0$):

$$\mathbf{C} = \begin{bmatrix} v_A & 0.0 & 0.0 \\ 0.0 & v_B + v_{(B,C)} & v_{(B,C)} \\ 0.0 & v_{(B,C)} & v_C + v_{(B,C)} \end{bmatrix} = \begin{bmatrix} 2.0 & 0.0 & 0.0 \\ 0.0 & 2.0 & 1.0 \\ 0.0 & 1.0 & 2.0 \end{bmatrix}$$

Observed covariance among species (N=100 simulations):

$$\overline{\mathbf{C}} = \begin{bmatrix} 1.85 & 0.11 & -0.24 \\ 0.11 & 2.08 & 0.84 \\ -0.24 & 0.84 & 1.94 \end{bmatrix}$$

from Revell et al. 2008



One can simply draw from this single distribution to 'simulate' BM evolution on a tree

from Revell et al. 2008

Brownian motion

- Brownian motion has been & continues to be an important model for studying the evolution of quantitative traits.
- It provides the theoretical basis for other methods that we will learn (such as independent contrasts, phylogenetic regression, and ancestral state reconstruction).
- It also provides the foundation for more complex models of trait evolution on trees.