

# Branch Lengths and Diversification

Lecture 3  
March 24, 2009

## Outline

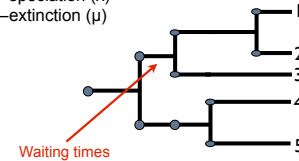
- How do lineages accumulate under the birth-death model?
- Estimating speciation and extinction rates from trees
- Testing hypotheses about diversification

## Outline

- How do lineages accumulate under the birth-death model?
- Estimating speciation and extinction rates from trees
- Testing hypotheses about diversification

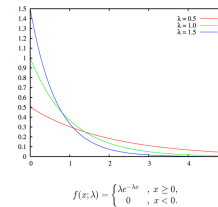
## Diversification Model

- Model: birth-death
- Every lineage has constant probability of:
  - speciation ( $\lambda$ )
  - extinction ( $\mu$ )



## Waiting times

- Under a birth-death model, the time intervals between successive events are always drawn from exponential distributions



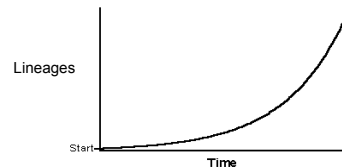
Expected species diversity under a birth-death model:

$$E[N_t] = N_0 e^{(b-d)t}$$

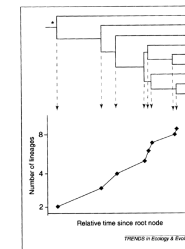
$N_t$  = species diversity after time  $t$   
 $N_0$  = starting species diversity  
 $b$  = birth rate  
 $d$  = death rate  
 $t$  = time

## Species Accumulation

- Expected number of species increases exponentially with  $\lambda - \mu$

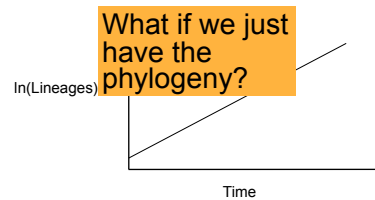


## Lineage-through-time

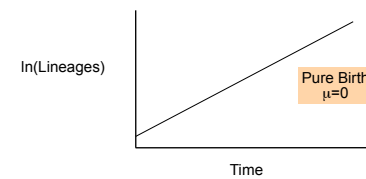


## Diversification Model

- Expected number of species increases exponentially with  $\lambda - \mu$

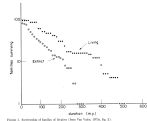


## LTT Plot



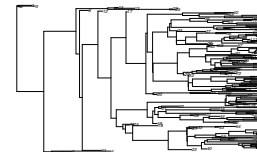
## Lineage-through-time plots

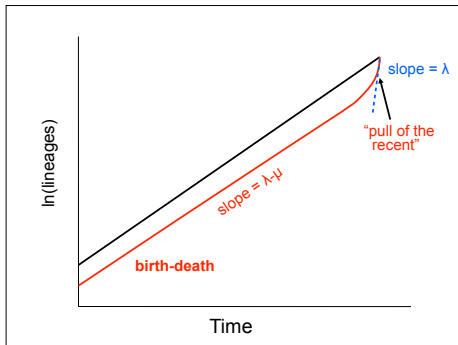
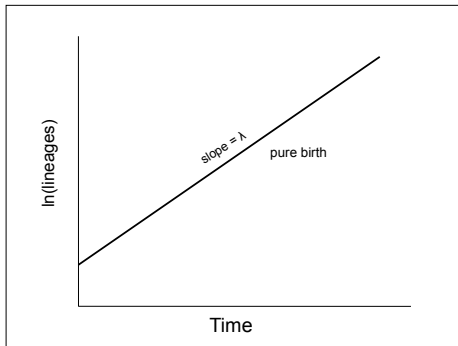
- Plot cumulative number of lineages through time
- Compare to expectations under null model
- Related to Van Valen's taxonomic survivorship curves



## Extinction

- Can leave an imprint on present-day phylogenies
- Older lineages are more likely to have gone extinct than younger ones





## Outline

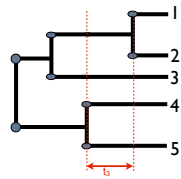
- How do lineages accumulate under the birth-death model?
- Estimating speciation and extinction rates from trees
- Testing hypotheses about diversification

## Pure Birth

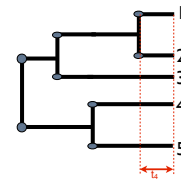
- Can use maximum likelihood to estimate parameters
- $L = \text{Pr}[\text{tree} \mid \lambda]$

$t_1$  = First waiting time  
 Drawn from exponential distribution with parameter  $2\lambda$   
 $\text{Pr}[t_1 = x] = 2\lambda e^{-2\lambda x}$

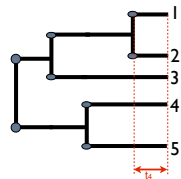
$t_2$  = Second waiting time  
 Drawn from exponential distribution with parameter  $3\lambda$   
 $\text{Pr}[t_2 = x] = 3\lambda e^{-3\lambda x}$



$t_3$  = Third waiting time  
 Drawn from exponential  
 distribution with parameter  $4\lambda$   
 $\Pr[t_3 = x] = 4\lambda e^{-4\lambda x}$



$t_4$  = Fourth waiting time (but no speciation!)  
 What is the probability of NO  
 speciation event in time  $t_4$ ?  
 Consider exponential distribution  
 with parameter  $5\lambda$   
 $\Pr[t_4 = x] = e^{-5\lambda x}$



$t_4$  = Fourth waiting time (but no speciation!)  
 What is the probability of NO  
 speciation event in time  $t_4$ ?  
 Consider exponential distribution  
 with parameter  $5\lambda$

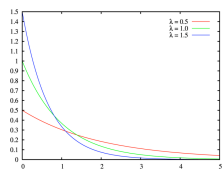
$$\begin{aligned}\Pr[t_1 = x] &= 2\lambda e^{-2\lambda x} \\ \Pr[t_2 = x] &= 3\lambda e^{-3\lambda x} \\ \Pr[t_3 = x] &= 4\lambda e^{-4\lambda x} \\ \Pr[t_4 = x] &= e^{-5\lambda x}\end{aligned}$$

$$\text{likelihood} = P[\text{tree} \mid \lambda] = 2\lambda e^{-2\lambda t_1} 3\lambda e^{-3\lambda t_2} 4\lambda e^{-4\lambda t_3} e^{-5\lambda t_4}$$

In general,

$$\text{likelihood} = P[\text{tree} \mid \lambda] = e^{-n\lambda s} \prod_{i=2}^{n-1} i\lambda e^{-i\lambda x_i} = (n-1)! \lambda^{n-2} e^{-\lambda s}$$

$$s = \sum_{i=2}^{n-1} i x_i$$



$$\begin{aligned}\Pr[\text{no event} \mid t=x] &= \\ 1 - \text{CDF} &= \\ 1 - (1 - e^{-\lambda x}) &= \\ e^{-\lambda x}\end{aligned}$$

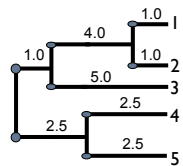
$$F(x, \lambda) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Cumulative distribution function

## Likelihood for a tree under a Pure-birth model

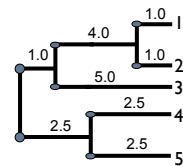
$$L = P[\text{tree} \mid \lambda] = (n-1)! \lambda^{n-2} e^{-\lambda s}$$

$n$  = number of taxa  
 $\lambda$  = speciation rate  
 $s$  = sum of all branch lengths in the tree



$$L = \Pr[\text{tree} \mid \lambda] = (n-1)! \lambda^{n-2} e^{-\lambda s}$$

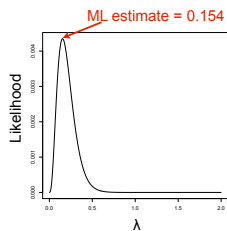
For  $\lambda = 1$ ,  
 $n = 5$   
 $s = 19.5$   
 $L = (5-1)! 1^3 e^{-19.5}$   
 $L = 8.16 \times 10^{-8}$



$$\hat{\lambda} = \frac{n-2}{s} = 0.154$$

$$SE_{\hat{\lambda}} = \frac{\hat{\lambda}^2}{n-2} = 0.0079$$

$0.154 \pm 0.0079$



## Birth-death models

- Can use maximum likelihood to estimate parameters of a birth-death process
- $L = \Pr[\text{tree} \mid \lambda, \mu]$
- Can be difficult to estimate  $\lambda$  and  $\mu$
- Easier to estimate composite parameters

$$r = \lambda - \mu \quad \epsilon = \mu/\lambda$$

## Analytic solution

$$\hat{\lambda} = \frac{n-2}{s}$$

$$SE_{\hat{\lambda}} = \frac{\hat{\lambda}^2}{n-2}$$

Maximum likelihood estimate of  $\lambda$  under PB model

