

O que é verossimilhança?



Sir Ronald Aylmer Fisher
(17 February 1890 – 29 July
1962)

$$L(\theta|x) = P(x|\theta)$$

A verossimilhança (L) de um conjunto de parâmetros (θ), dado alguma observação (x) é igual a probabilidade daquela observação ter ocorrido dados os valores daqueles parâmetros.

Parâmetros estatísticos podem ser vistos como as características numéricas de um modelo.

Um **modelo** estatístico é a formalização matemática da relação entre variáveis que correspondem a observações potenciais que inclui a descrição das incertezas sobre estas observações devido à variabilidade natural, erros ou informação incompleta.

O que é verossimilhança?



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Fisher (1912) empregou a função de verossimilhança $f(\theta|x)$ com a ideia de que o(s) valor(es) de θ que maximizam a probabilidade dos dados observados (x) seria um bom estimador de θ .

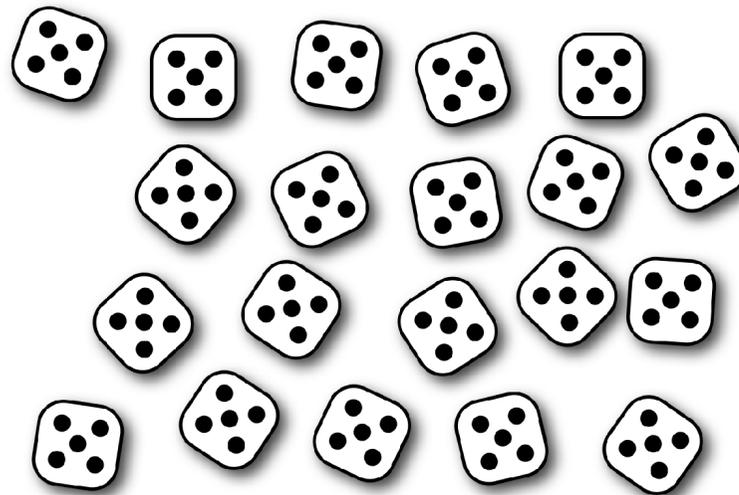
O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$

Considere:

Você jogou 20 dados sobre uma mesa e todos apresentaram 5!
Portanto:

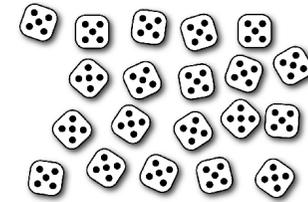
$x =$



O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$

$$x =$$



Considere que você deseja avaliar a L de dois modelos (θ_1 e θ_2):

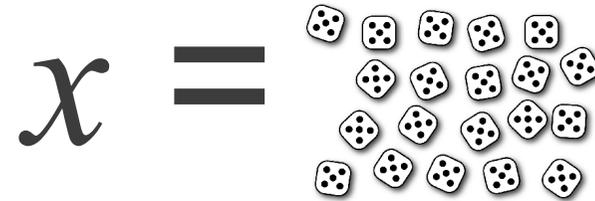
Modelo 1 (θ_1) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

Modelo 2 (θ_2) – dado viciado – assume que todas as faces do dado tem o número 5. Portanto, a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1)^{20}$.

Um **modelo** estatístico é a formalização matemática da relação entre variáveis que correspondem a observações potenciais que inclui a descrição das incertezas sobre estas observações devido à variabilidade natural, erros ou informação incompleta.

O que é verossimilhança?

$$L(\theta|x) = P(x|\theta)$$



Considere que você deseja avaliar a L de dois modelos (θ_1 e θ_2):

Modelo 1 (θ_1) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

$$L(\theta_1|x) = P_{(x|\theta_1)} = (1/6)^{20} = 1/3.656.158.440.062.976$$

Modelo 2 (θ_2) – dado viciado – assume que todas as faces do dado tem o número 5. Portanto, a probabilidade conjunta (i.e., $P_{(A \text{ e } B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1)^{20}$.

$$L(\theta_2|x) = P_{(x|\theta_2)} = (1)^{20} = 1$$

Verossimilhança vs. Probabilidade

$$x = \text{🎲🎲🎲🎲🎲} = \text{observação} = O$$

Modelo 1 (θ_1) – dado honesto = Hipótese 1 = H_1

Modelo 2 (θ_2) – dado viciado = Hipótese 2 = H_2

$$P(O|H) \neq P(H|O)$$

Verossimilhança vs. Probabilidade

$$P(O|H) \neq P(H|O)$$

Considere:

O = Você está ouvindo um barulho no forro de sua casa.

H = há *gremlins* jogando boliche no seu forro.



$$L_{(H|O)} = P_{(O|H)} = \text{Alta}$$

A verossimilhança desta hipótese é muito alta, pois se há *gremlins* jogando boliche no seu forro, a probabilidade de haver barulho é alta.

$$P_{(H|O)} = \text{Baixa}$$

No entanto, certamente você não pensa que o barulho é evidência (torna-se provável) que haja *gremlins* jogando boliche no seu forro.

Voltemos a ideia original:



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Fisher (1912) empregou a função de verossimilhança $f(\theta|x)$ com a ideia de que o(s) valor(es) de θ que maximizam a probabilidade dos dados observados (x) seria um bom estimador de θ .

Estimativa de máxima verossimilhança:

Descendente **CCACGTACGTACGTACGTACGTACGTACGTACGTAC**



$$\theta = \alpha t = v$$

onde v é a "distância evolutiva entre essas sequências."

Ancestral **TTACGTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

Estimativa de máxima verossimilhança:

Descendente **CCACGTACGTACGTACGTACGTACGTACGTAC**



$$\theta = \alpha t = v$$

onde v é a "distância evolutiva entre essas sequências."

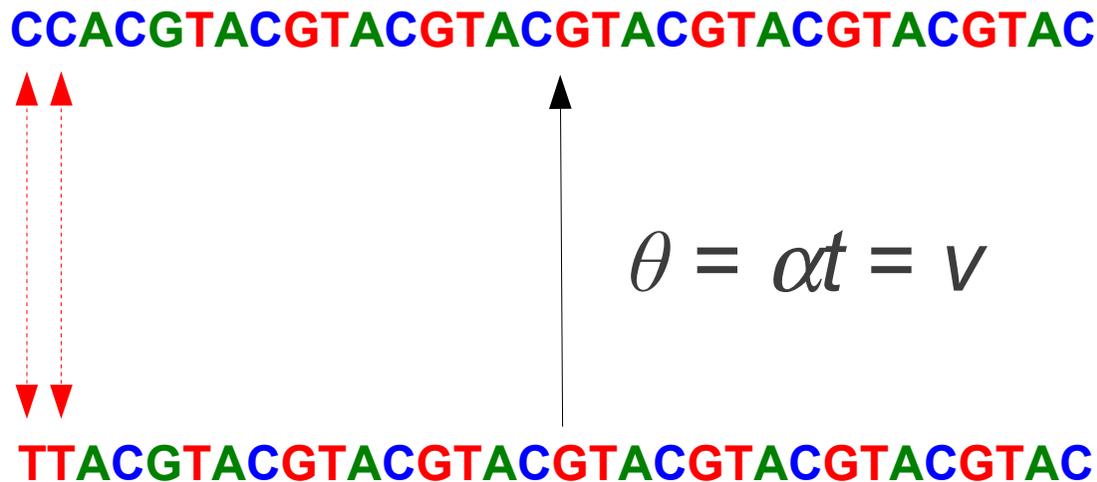
Ancestral **TTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

Estimativa de máxima verossimilhança:



probabilidade de não mudança de estado*

$$P(ii) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]$$

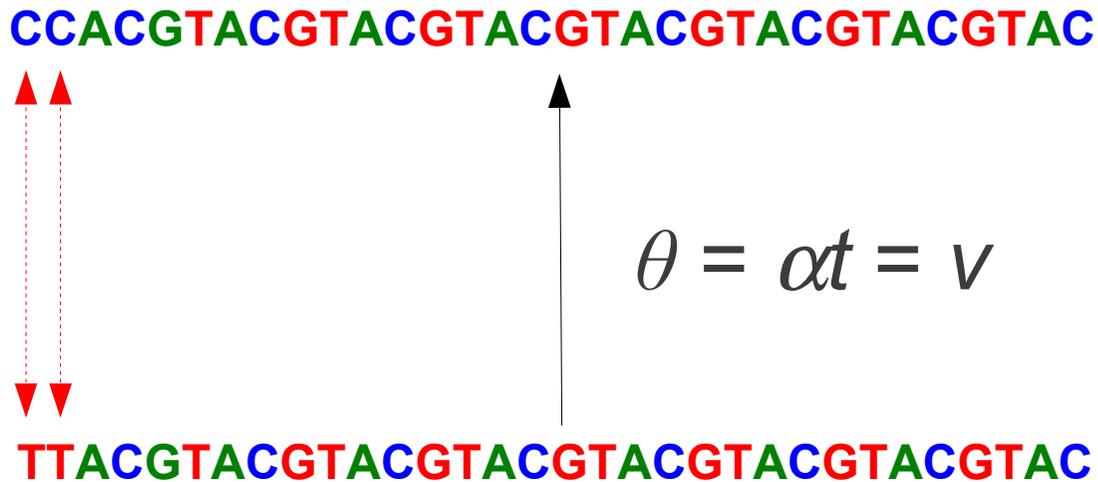
probabilidade de mudança de estado*

$$P(ij) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]$$

* baseado no modelo de JC69 (distribuição de Poisson) e frequências iguais de pares de base.

Voltemos a ideia original:

Estimativa de máxima verossimilhança:



probabilidade de não
mudança de estado*

$$P(ii) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4\nu} \right) \right]$$

probabilidade de
mudança de estado*

$$P(ij) = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4\nu} \right) \right]$$

$$L = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4\nu} \right) \right]^2 * \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4\nu} \right) \right]^{30}$$

* baseado no modelo de JC69 (distribuição de Poisson) e frequências iguais de pares de base.

Voltemos a ideia original:



Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

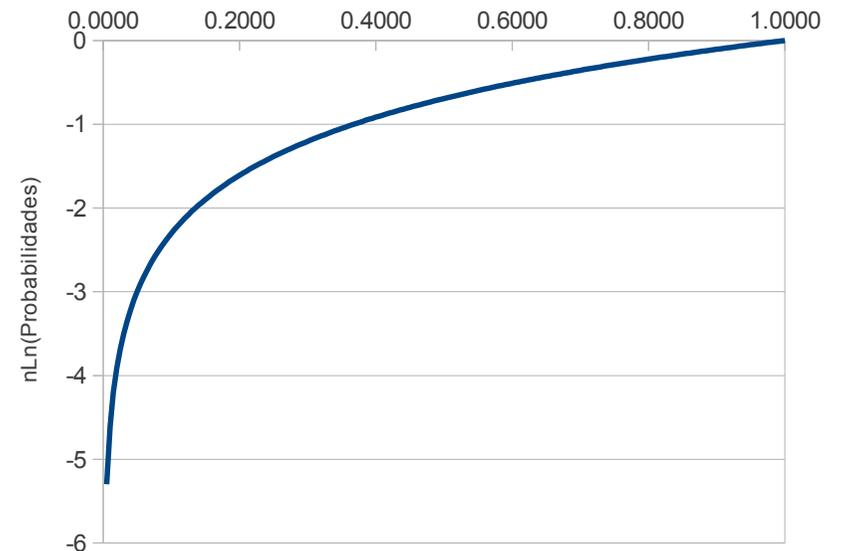
$$L = \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} - \frac{1}{4} e^{-4v} \right) \right]^2 * \left[\left(\frac{1}{4} \right) * \left(\frac{1}{4} + \frac{3}{4} e^{-4v} \right) \right]^{30}$$

Estimativa de Máxima Verossimilhança (v)



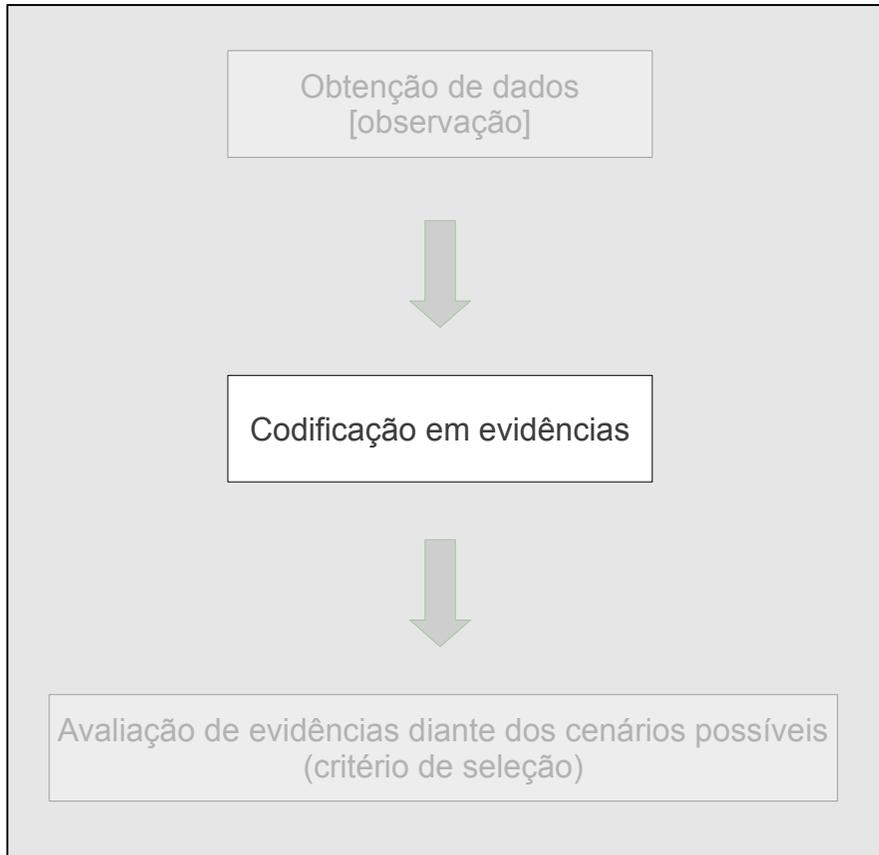
Estimativa de distância evolutiva (v)

nLn(Probabilidades) vs. Probabilidades



Probabilidades

Lógica da inferência filogenética



↓ ↓ ↓
 sp.X CTGGCTACGT
 sp.A TGGAGTAAGT
 sp.B CCTAGCAAGT
 sp.C CCTGATTGCA

Parcimônia:

EVIDÊNCIAS: transformação de estados de caracteres

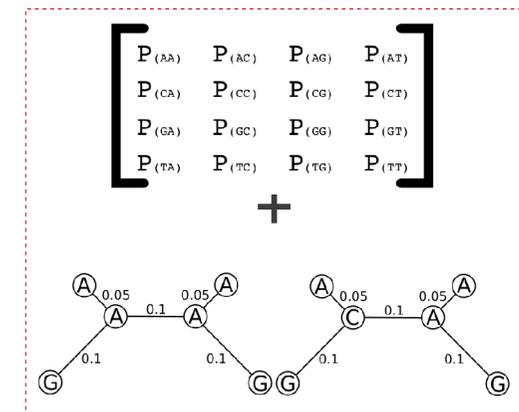
1a. posição: C ↔ T
 3a. posição: G ↔ T
 10a. posição: T ↔ A

CODIFICAÇÃO: matriz de dados

	c ₁	c ₂	c ₃	c ₄	c ₅	c ₆	c ₇	c ₈	c ₉	c ₁₀
sp.X	1	3	2	2	1	3	0	1	2	3
sp.A	3	2	2	0	2	3	0	0	2	3
sp.B	1	1	3	0	2	1	0	0	2	3
sp.C	1	1	3	2	0	3	3	2	1	0

Probabilística (ML):

EVIDÊNCIAS: modelo de transformações + topologia que melhor explicam seus dados.



sp.X CTGGCTACGT
 sp.A TGGAGTAAGT
 sp.B CCTAGCAAGT
 sp.C CCTGATTGCA

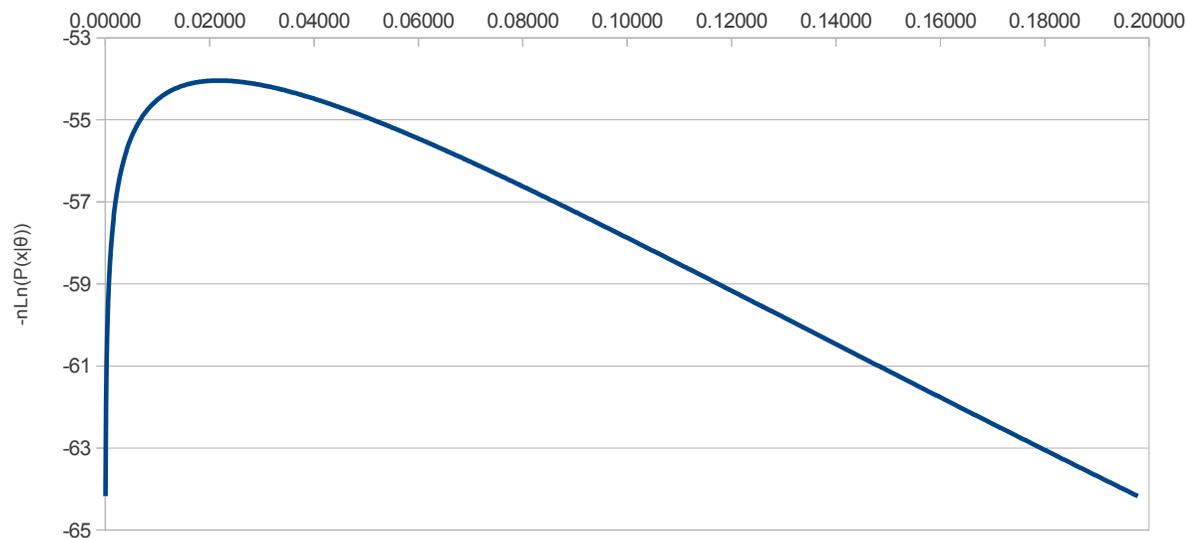
Função objetiva: Likelihood

$$L(\theta|x) = P(x|\theta)$$

Qual o valor de θ que maximiza a probabilidade dos dados observados?

$$L = \Pr \left(\begin{array}{l} \text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \end{array} \middle| \begin{array}{l} \begin{bmatrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{bmatrix} + \begin{array}{l} \text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \end{array} \right)$$

Estimativa de Máxima Verossimilhança (v)



Estimativa de distância evolutiva (v)

Probabilidades:

Modelo probabilístico: Representação matemática de um fenômeno aleatório.

Se dois eventos não possuem soluções em comum, eles são chamados de **disjuntos** e suas probabilidades obedecem à seguinte regra:



Regra 1: Se dois eventos A e B são disjuntos, então a probabilidade de qualquer um destes eventos e a soma das probabilidades dos dois eventos: $P_{(A \text{ ou } B)} = P_{(a)} + P_{(b)}$

Regra 2: A probabilidade de um evento **A** não ocorrer é $P_{(A^c)} = 1 - P_{(A)}$



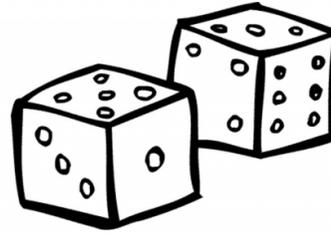
Copyright © Ron Leishman

$$P_{(\text{1})} = A/S = 1/6$$

$$P_{(\text{1})}^c = 1 - P_{(A)} = 1 - 1/6$$

Probabilidades:

Considere que você jogue dois dados:



A probabilidade de obter “olhos de serpente” é expressa pela seguinte regra:

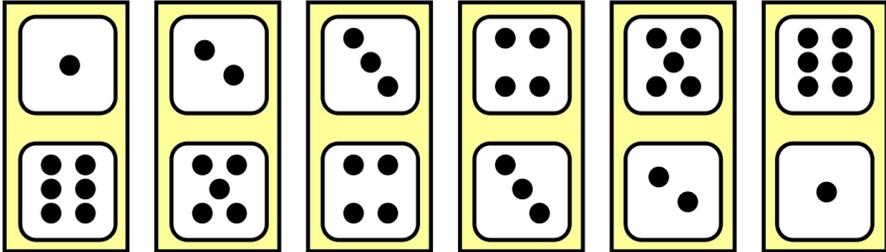
Regra 3: Se dois eventos A e B são independentes, então a probabilidade dos dois eventos ocorrerem um após o outro (i.e., intersecção de eventos, ou **probabilidade conjunta**) é o produto das probabilidades de cada evento: $P_{(A \text{ e } B)} = P_{(A)} * P_{(B)}$.

Combinando as regras:

Qual a probabilidade de se obter a soma 7 ao jogar dois dados?

E

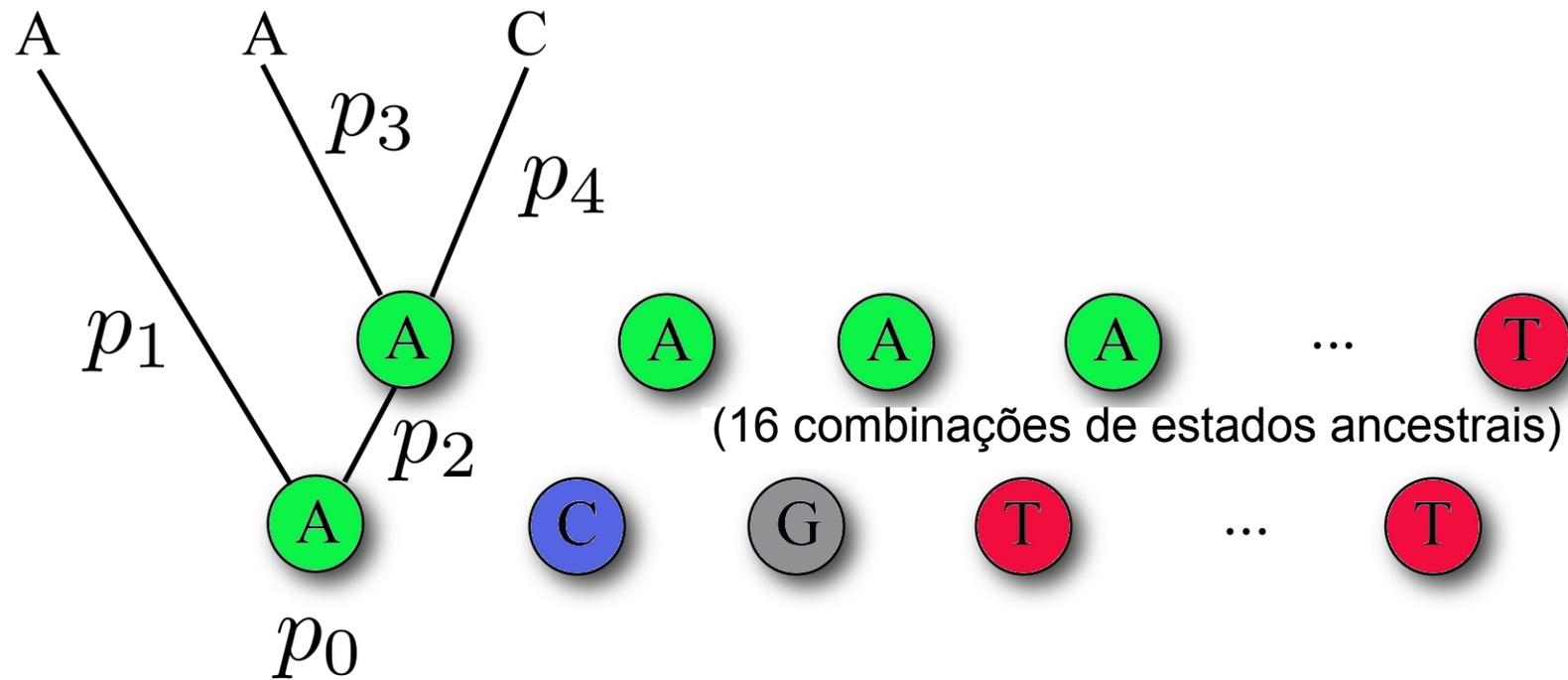
$1/6 * 1/6$

Six pairs of dice are shown, each pair enclosed in a yellow rectangular box. The pairs represent the following combinations of faces: (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), and (6, 1).

$(1/36)+(1/36)+(1/36)+(1/36)+(1/36)+(1/36) = 1/6$

OU OU OU OU OU

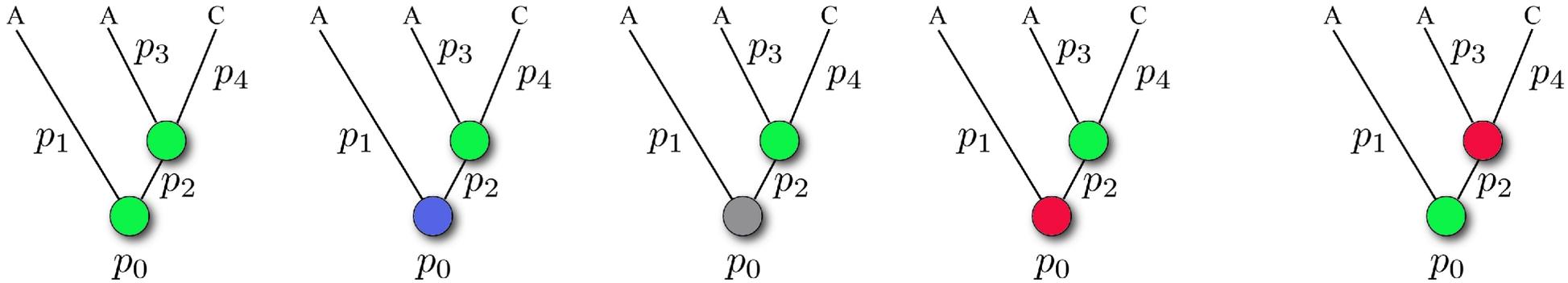
Lógica da inferência filogenética: Likelihood (MAL)



Regra 5: lintersecção de eventos, ou probabilidade conjunta: $P_{(A \text{ e } B)} = P_{(A)} * P_{(B)}$.

$$\text{Reconstrução 1} = P_0 * P_1 * P_2 * P_3 * P_4$$

Lógica da inferência filogenética: Likelihood



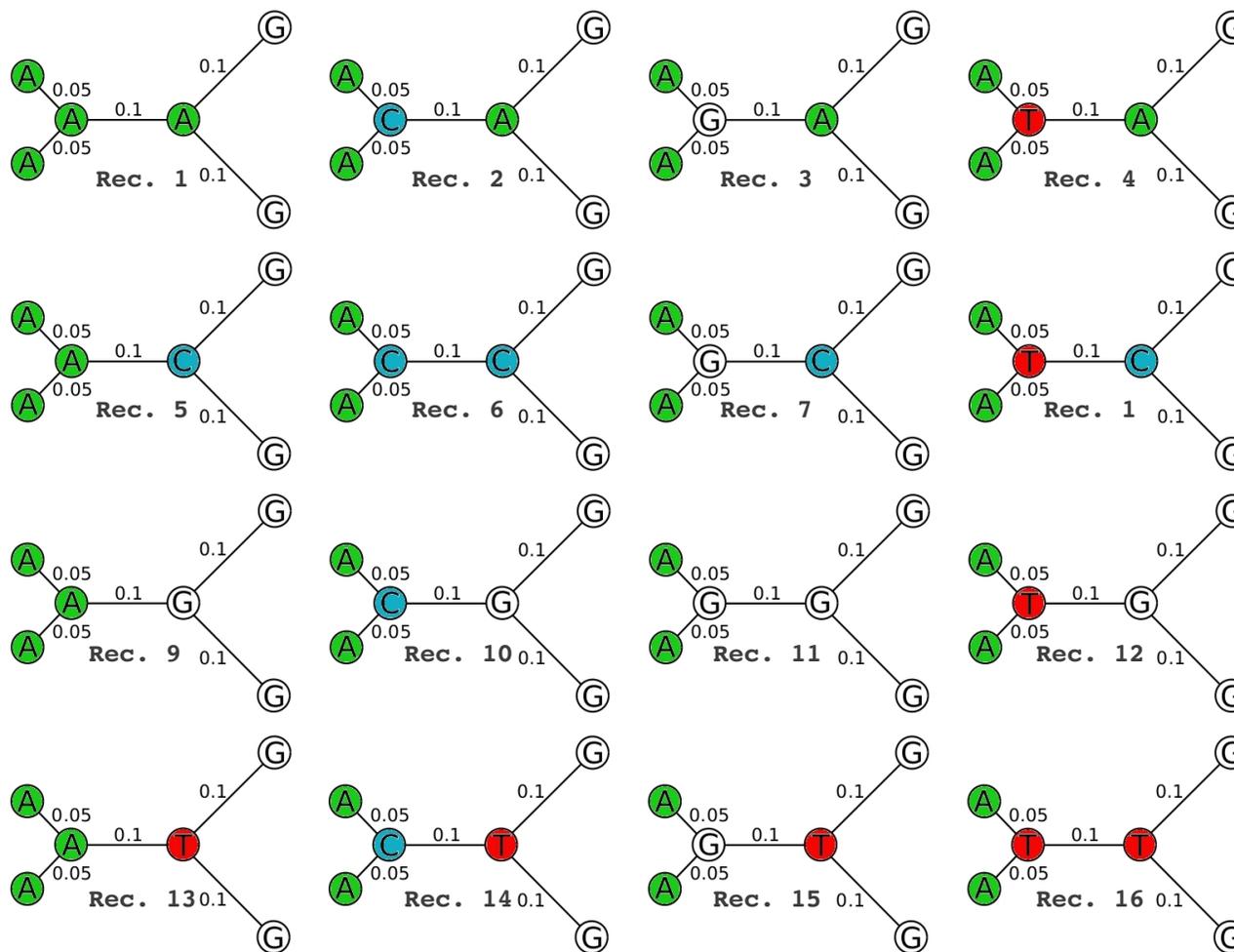
Reconstrução 1 + Reconstrução 2 + Reconstrução 3 + Reconstrução 4 + ... + Reconstrução 16

Regra 3: Eventos disjuntos: $P_{(A \text{ ou } B)} = P_{(a)} + P_{(b)}$

Lógica da inferência filogenética: Likelihood

$$L = \Pr \left(\begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix} \mid \begin{bmatrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{bmatrix} + \begin{matrix} \text{sp. X} & \text{CTGGCTACGT} \\ \text{sp. A} & \text{TGGAGTAAGT} \\ \text{sp. B} & \text{CCTAGCAAGT} \\ \text{sp. C} & \text{CCTGATTGCA} \end{matrix} \right)$$

Avaliação de todas as reconstruções possíveis

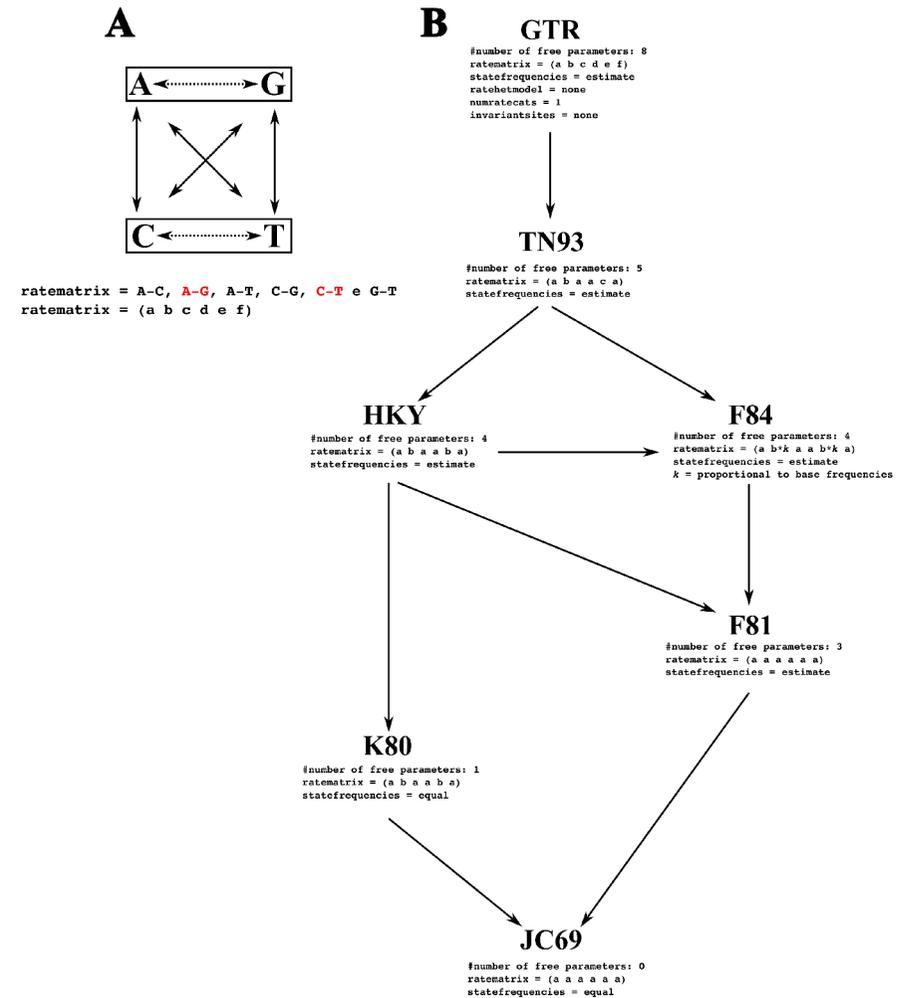


Lógica da inferência filogenética: Likelihood

$$L = \Pr \left(\begin{array}{l} \text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \end{array} \mid \begin{bmatrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{bmatrix} + \begin{array}{c} \text{sp. X CTGGCTACGT} \\ \text{sp. A TGGAGTAAGT} \\ \text{sp. B CCTAGCAAGT} \\ \text{sp. C CCTGATTGCA} \end{array} \right)$$

Avaliação de todas as reconstruções possíveis

$$Q = \begin{bmatrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{bmatrix}$$



Lógica da inferência filogenética: o cálculo

Para,	1	2..	j		N
sp.X	C	TGG	C	T	A...CGT
sp.A	T	GGAG	T	A...AGT	
sp.B	C	TAG	C	A...AGT	
sp.C	C	TGAT	T	A...GCA	

A Verossimilhança Máxima de um determinado caráter (i.e., sítio) é:

$$L_{(j)} = P_{(rec. 1)} + P_{(rec. 2)} + P_{(rec. 3)} + \dots + P_{(rec. n)}$$

A Verossimilhança Máxima (L^*) de uma determinada hipótese é dada por:

$$L = L_{(1)} * L_{(2)} * L_{(3)} * \dots * L_{(n)} = \prod_{j=1}^N L_{(j)}$$

* tradicionalmente ela é avaliada pela soma dos logaritmos neperianos das probabilidades de cada caráter:

$$\ln L = \ln L_{(1)} + \ln L_{(2)} + \ln L_{(3)} + \dots + \ln L_{(n)} = \sum_{j=1}^N \ln L_{(j)}$$

Lógica da inferência filogenética: modelos complexos

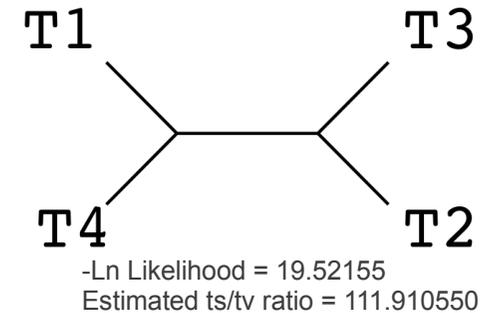
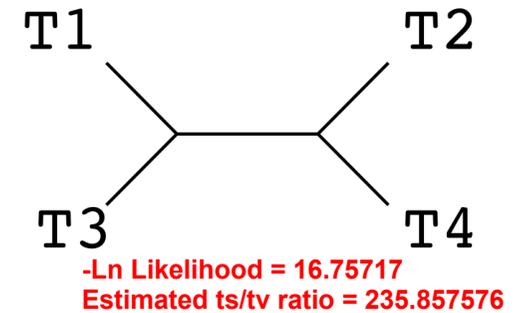
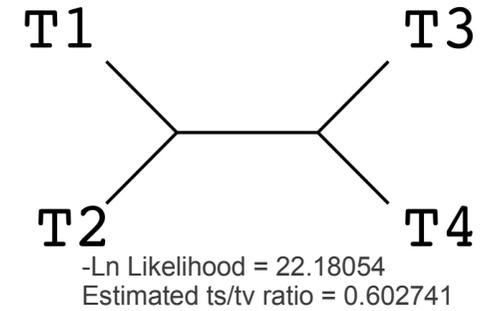
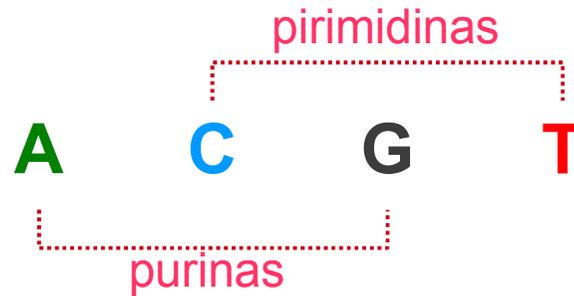
K2P

Kimura Two Parameters (K80)

Considere:

T1 A C G T
 T2 C G T A
 T3 G T A C
 T4 T A C G

	A	C	G	T
A	$P_{(A,A)}$	$P_{(A,C)}$	$P_{(A,G)} * k$	$P_{(A,T)}$
C	$P_{(C,A)}$	$P_{(C,C)}$	$P_{(C,G)}$	$P_{(C,T)} * k$
G	$P_{(G,A)} * k$	$P_{(G,C)}$	$P_{(G,G)}$	$P_{(G,T)}$
T	$P_{(T,A)}$	$P_{(T,C)} * k$	$P_{(T,G)}$	$P_{(T,T)}$



... onde k é a razão entre eventos de transição e transversão (valor estimado neste exemplo).

Note que utilizando o critério de parcimônia, nenhum caráter seria considerado informativo!

Likelihood vs. Probabilidade:

Likelihood Criterion:

$$L_{(\text{mod.} | \text{obs.})} = \text{Pr} (\text{observação} | \text{modelo})$$

Bayesian Criterion:

$$P_{(H|O)} = \frac{P_{(O|H)} P_{(H)}}{P_{(O)}}$$

Likelihood

prior probability

probabilidade incondicional da observação

Escolha de critérios de otimização

Qual é o melhor método?

Qual é a melhor árvore (topologia)?

Não há uma única resposta para a pergunta e para ambas perguntas é necessário adotar um critério para avaliação objetivo!

MÉTODOS DE DISTÂNCIA E DE TRANSFORMAÇÃO DE CARACTERES

Problemas centrais com o métodos de distância:

Impossibilidade de identificar transformações

Impossibilidade de reconstruir estados ancestrais hipotéticos

Impossibilidade de postular hipóteses de homologia

Escolha de critérios de otimização

Parâmetros de avaliação:

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Escolha de critérios de otimização

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

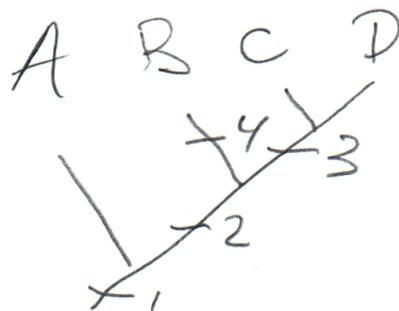
Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

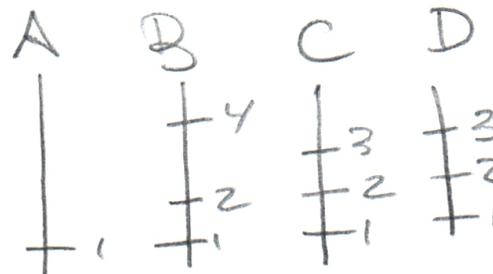
1. Navalha de Ockham e argumentação Popperiana

"The minimization approach of parsimony offers a specific method to achieve the base-line goal of choosing genealogical explanation of variation, and of identifying non-conforming observations that require additional modes of causality."

(Wheeler 2012:271)



ou



Escolha de critérios de otimização

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

2. Parcimônia e processos evolutivos

procedimento de inferência

vs.

Modelo de evolução

- i. Não assume que evolução ocorre parcimoniamamente (ex. Homoplasia)
- ii. Em ciência utilizada como um processo de descobrimento e teste
- lii. Pode ser defendida sem apelo ao autoritarismo a apriorismo

Escolha de critérios de otimização

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)

Como nós "sabemos" as coisas?

Como hipóteses científicas são propostas e testadas?

Qual é a relação entre observação e inferência?

3. Indução e estimativa estatística

Sistemática, assim como muitas outras disciplinas em Biologia, é uma ciência histórica na qual os eventos que buscamos explicar são únicos no tempo e no espaço.

"If each character observation is a unique object, it cannot be a random sample drawn from a parameterized distribution expressed as a model of change (e.g., GTR+I+G)."

"By this rationale, parsimonious trees have explanatory power in minimizing ad hoc hypotheses, and increased observation will increase the severity of test, but would not reduce any measure of 'sampling error', since there is no sample distribution that is measured."

(Wheeler 2012:272)

Escolha de critérios de otimização

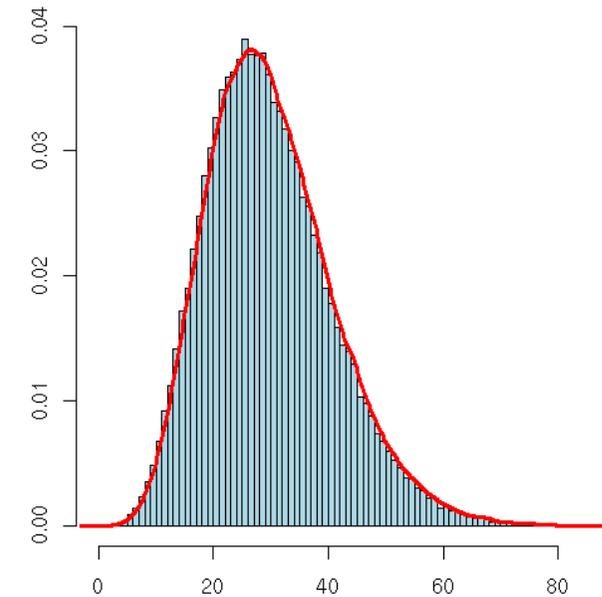
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Bases analíticas: Consistência, Eficiência e Robustez

Bases empíricas:

Em sua grande maioria baseada em dados restritos ou simulados.

Objeção: se as condições sobre os quais os resultados foram gerados ou provados se aplicam a casos empíricos (i.e., dados reais).

Escolha de critérios de otimização

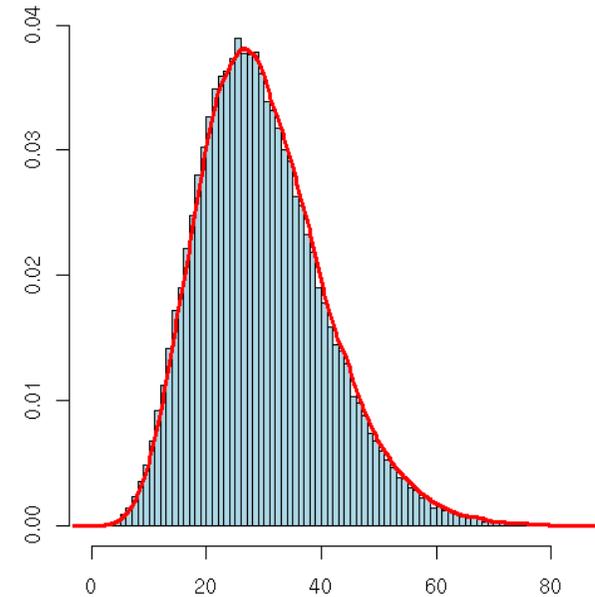
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades: duas interpretações

i. Grau de crença:

Confiança relativa em determinado resultado, considere

"Esta moeda tem 80% de chance de ser honesta."

Mas a moeda é ou não é honesta!



Escolha de critérios de otimização

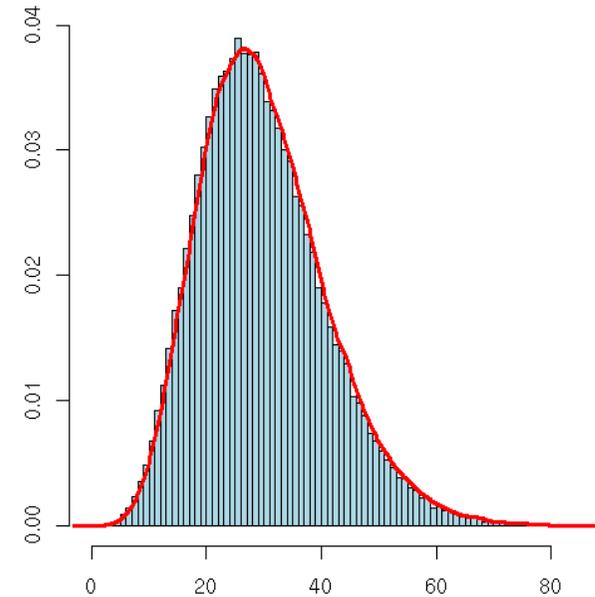
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades: duas interpretações

ii. Uma proposição verdadeira sobre o universo natural semelhante à eventos repetidos:

Implica em uma proposição sobre o que é ontologicamente real, considere:

"Esta moeda tem 80% de chance de dar coroa em uma jogada."



Oferece uma propriedade da moeda e uma proposição relacionada à sua realidade física.

Escolha de critérios de otimização

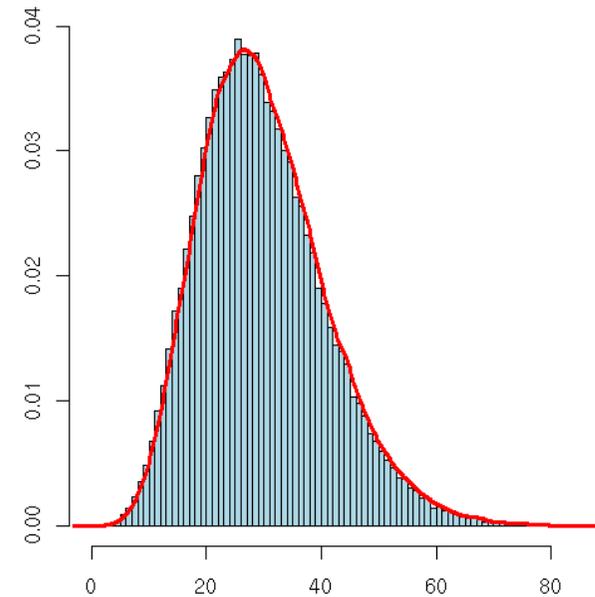
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"Qual é a probabilidade de que Tiradentes e Joaquim José da Silva Xavier eram a mesma pessoa?"

Escolha de critérios de otimização

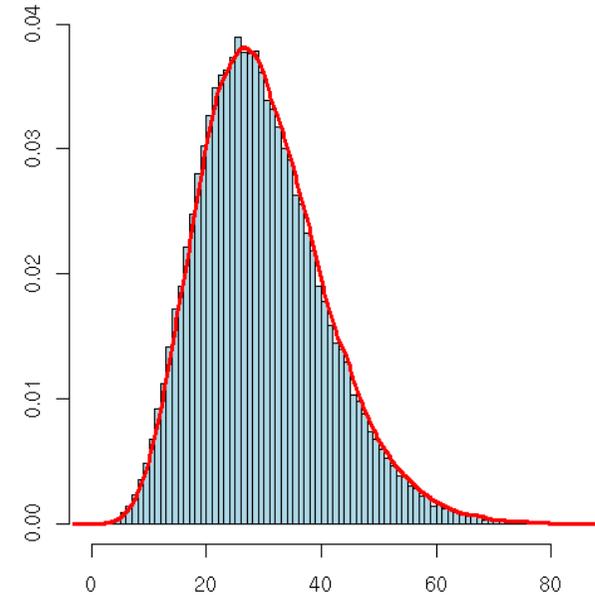
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"A e B são grupos-irmãos?"

Escolha de critérios de otimização

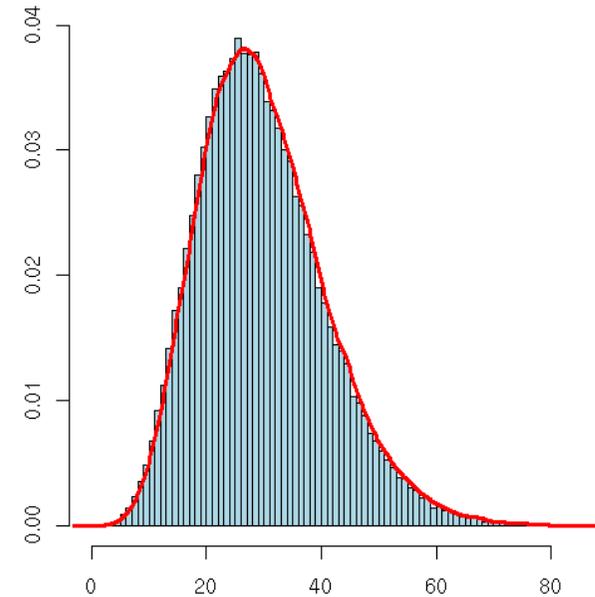
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Considere:

"Asas de morcego e asas de passarinho são homólogas?"

After this short introduction to Bayesian methods, one may wonder what the appeal is over maximum likelihood (ML) and parsimony based methods in phylogenetic analysis. There are both theoretical and practical advantages. Arguably, the conflict between probabilistic based methods and cladistic parsimony as chronicled by Felsenstein (2001) has resulted in a split into two different schools where the probabilists use model based methods and the others do not. Without models it is hard to quantify the uncertainty in an analysis, which is a requisite of doing science.

Escolha de critérios de otimização

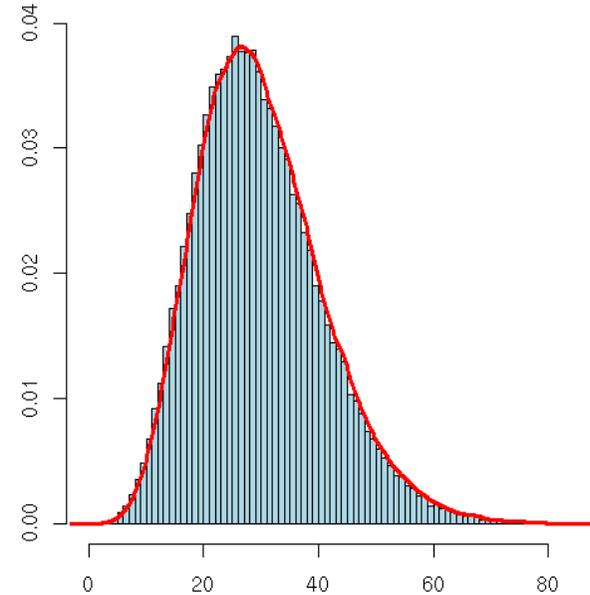
Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

Performance

Congruência (?)



Probabilidades:

Ciências históricas não permitem facilmente a interpretação frequencista (i.e., ii)

Nesse contexto, qualquer interpretação de probabilidade está restrita ao nosso grau de crença. Portanto probabilidades associadas a eventos históricos conferem:

- i.* grau de suporte entre hipóteses e não sua realidade;
- ii.* elas não descrevem o universo natural, mas sim nosso entendimento dele.

"We cannot say whether a given set of historical statements (i.e., a tree) is true, but we can say we have relative degree of belief based on empirical observation and an optimality criterion"

(Wheeler 2012:274)

Escolha de critérios de otimização

Méritos relativos dos métodos:

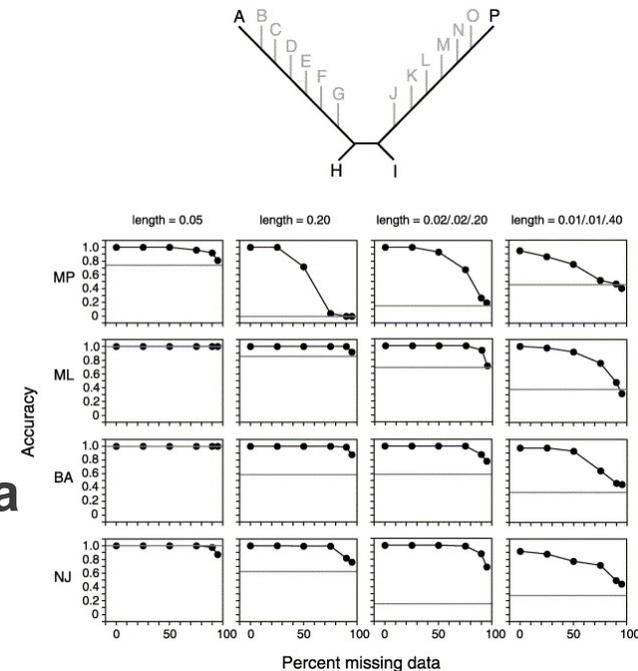
Epistemologia

Comportamento estatístico

Performance

Congruência (?)

**Requer o conhecimento da verdadeira topologia
primariamente restrita a dados simulados**



Long Branch Attraction (LBA - Felsenstein Zone):

Principal crítica à parcimônia, mas o mesmo se aplica a métodos de distância e a CMMRL

MRL --> Farris Zone ou Anti Felsenstein Zone ou Long Branch Repulsion

Todos os métodos, dado algumas circunstâncias apresentam o mesmo comportamento.

Congruência:

Considerado o mais poderoso teste para hipóteses filogenéticas

Poucos consideram esse parâmetro (e.g., Wheeler, 2006)

Escolha de critérios de otimização

Como escolher?

Epistemologia ou consistência estatística não oferecem caminho inequívoco

- i. Todos os métodos podem participar do processo hipotético-dedutivo.
- ii. Todos estão sujeitos a comportamentos inconsistentes.

Uma possibilidade: sincretismo

- i. Convergência de resultados é trivial, embora rara.
- ii. Na maioria dos casos há incongruências, o que fazer?
 - a. critérios de otimização possuem diferentes motivações e epistemologias subjacentes.
 - b. Consenso é um caminho ilógico.

Escolha de critérios de otimização

Como escolher?

Critérios de otimização são premissas analíticas utilizadas para interpretar eventos historicamente únicos. Como tal, eles não podem ser testados empiricamente por acurácia. Por outro lado, congruência pode ser um caminho a ser seguido. No entanto, em seu cerne, critérios de otimização são pressupostos que requer suporte e justificativa quando empregados. Pressupostos diferentes levam a resultados diferentes. Somente justificando nossas premissas nós podemos defender nossas conclusões.