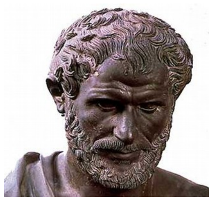
Conceitos fundamentais:

Comprimento de ramo
Estimativas de tempo absoluto
Suporte e resolução
Descrição de relações de parentesco
Notação parentética
Diagramas (árvores) de consenso
- *Cosenso estrito*
- *Consenso de maioria*

O desenvolvimento da Cladística

Dados moleculares e modelos probabilísticos

Aristóteles – 384-322 A.C.



Darwin

1809-1882

1859

Período essencialista

Mundo dinâmico

Resistência e Nova Síntese

Sistemática Evolutiva

1936 - 1947

1960's

Fenética

1970's

Cladística

1990's

Probabilismo

Carolus Linnaeus

1707-1778

Buffon

1707-1788

Lamarck

1744 -1829

St-Hilair

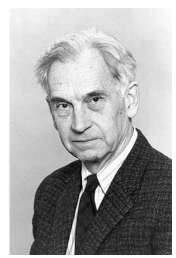
1772 -1844

Cuvier

1769 -1832

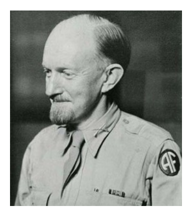
Ernest Mayr

1904 - 2005



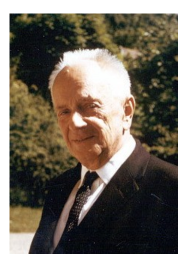
G.G. Simpson

1902 - 1984

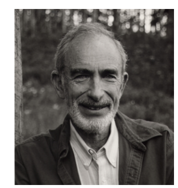


Theodosius Dobzhansky

1900 -1975



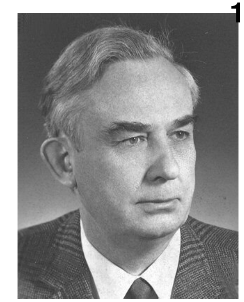
Paul Erlich



James Rohlf

R. Sokal

1926 -

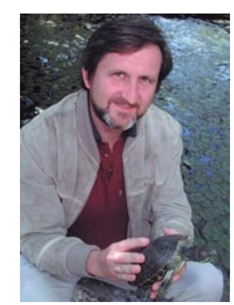


Willi Hennig

1913 - 1976

Steve Farris

Joe Felsenstein



David Hillis

Motivação pára o uso de Verossimilhança:

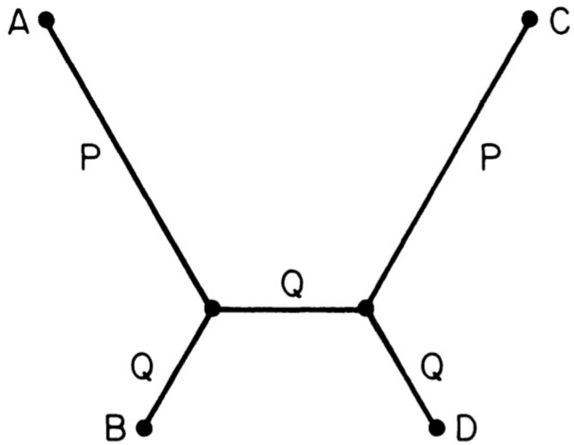
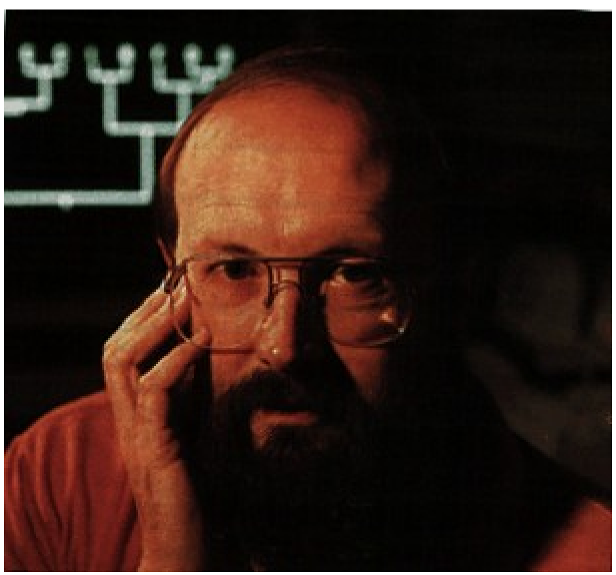


FIG. 3.—True unknown phylogeny (with root omitted) used to find cases in which unrooted Wagner tree parsimony methods will be inconsistent.



Joe Felsenstein
University of
Washington

CASES IN WHICH PARSIMONY OR COMPATIBILITY METHODS WILL BE POSITIVELY MISLEADING[1]

JOSEPH FELSENSTEIN

Abstract

Felsenstein, J. (Department of Genetics, University of Washington, Seattle, WA 98195) 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.—For some simple three- and four-species cases involving a character with two states, it is determined under what conditions several methods of phylogenetic inference will fail to converge to the true phylogeny as more and more data are accumulated. The methods are the Camin-Sokal parsimony method, the compatibility method, and Farris's unrooted Wagner tree parsimony method. In all cases the conditions for this failure (which is the failure to be statistically consistent) are essentially that parallel changes exceed informative, nonparallel changes. It is possible for these methods to be inconsistent even when change is improbable a priori, provided that evolutionary rates in different lineages are sufficiently unequal. It is by extension of this approach that we may provide a sound methodology for evaluating methods of phylogenetic inference. [Numerical cladistics; phylogenetic inference; maximum likelihood estimation; parsimony; compatibility.]

Parsimony or minimum evolution methods were first introduced into phylogenetic inference by Camin and Sokal (1965). This class of methods for inferring an evolutionary tree from discrete-character data involves making a reconstruction of the changes in a given set of characters on a given tree, counting the smallest number of times that a given kind of event need have happened, and using this as the measure of the adequacy of the evolutionary tree. (Alternatively, one can compute the weighted sum of the numbers of times several different kinds of events have occurred.) One attempts to find that evolutionary tree which requires the fewest of these evolutionary events to explain the observed data. Camin and Sokal treated the case of irreversible changes along a character state tree, minimizing the number of changes of character states required. A number of other parsimony methods have since appeared in the systematic literature (Kluge and Farris, 1969; Farris, 1969, 1970, 1972, 1977; Farris, Kluge, and Eckhardt, 1970) and parsimony methods have also found widespread use in studies of molecular evolution (Fitch and Margoliash, 1967, 1970; Dayhoff and Eck, 1968; see also Fitch, 1973). Cavalli-Sforza and Edwards (1967; Edwards and Cavalli-Sforza, 1964) earlier formulated a minimum evolution method for continuous-character data.

An alternative methodology for phylogenetic inference is the compatibility method, introduced by Le Quesne (1969, 1972). He suggested that phylogenetic inference be based on finding the largest possible set of characters which could simultaneously have all states be uniquely derived on the same tree. The estimate of the phylogeny is then taken to be that tree. While Le Quesne's specific suggestions as to how this might be done have been criticized by Farris (1969), his general approach, which is based on Camin and Sokal's (1965) concept of the compatibility of two characters, has been made rigorous and extended in a series of papers by G. F. Estabrook, C. S. Johnson, Jr., and F. R. McMorris (Estabrook,

[1] This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

Motivação pára o uso de Verossimilhança:

"Felsenstein Zone" Trees

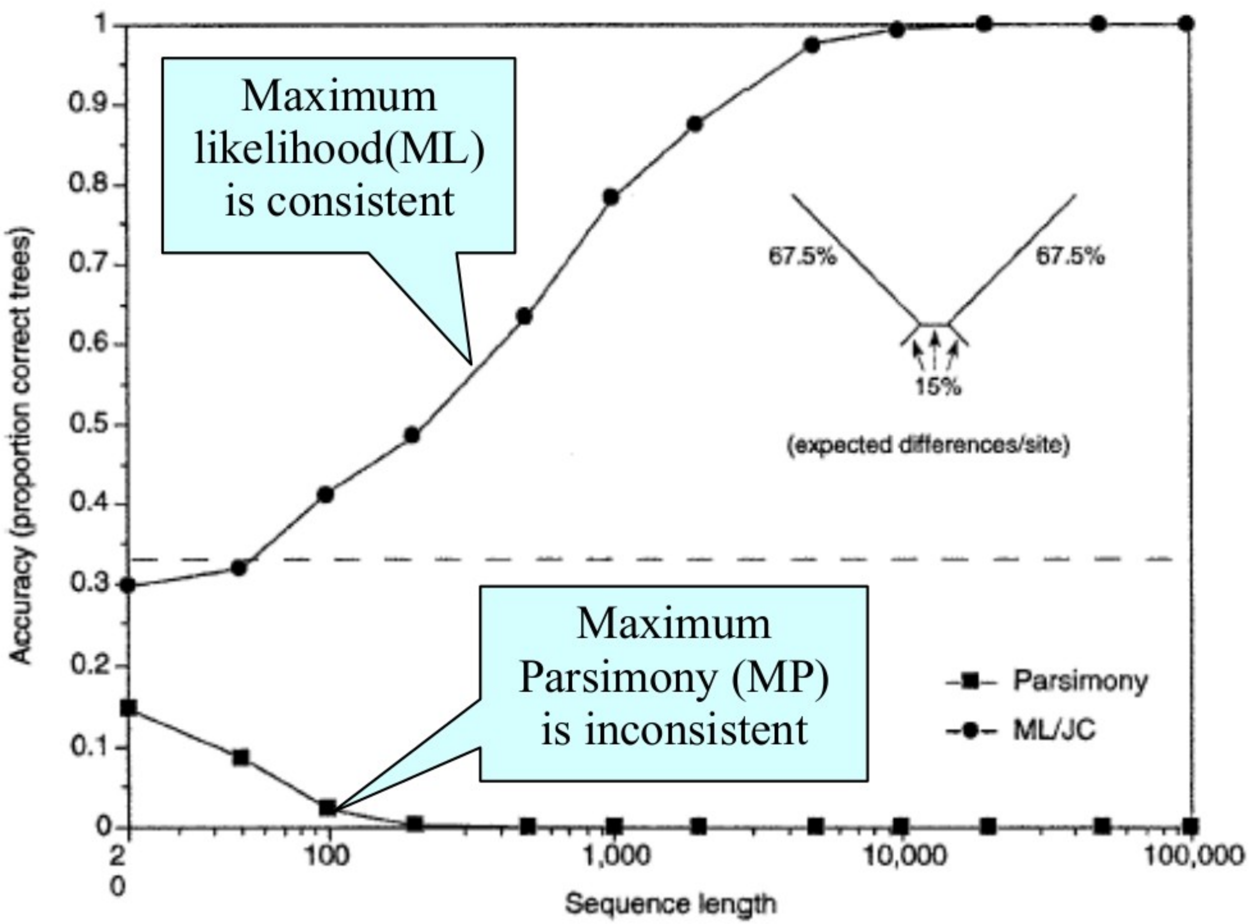


Fig. 3, p. 531, from Swofford et al. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology* 50: 525-539.

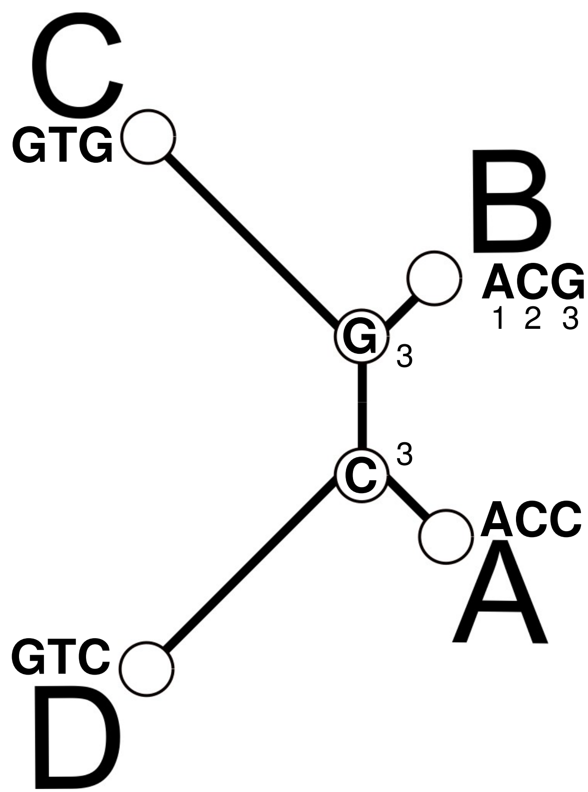
Accuracy is proportion correct out of 1000 simulation replicates

Dashed line is 1/3, the probability of getting the tree correct by pure random chance

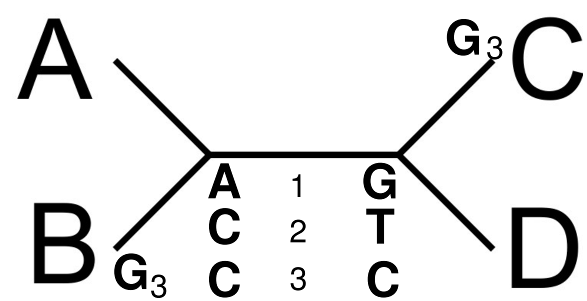
Fonte: Paul O. Lewis (UCONN), 2014

Motivação pára o uso de Verossimilhança:

ML



MP



O que é verossimilhança?

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

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

A verossimilhança (L) de um conjunto de parâmetros (θ), dado alguma observação (x) é proporcional 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 February 1890 – 29 July 1962)

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

$$f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

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.

Estimativa de parâmetros:

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

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

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.

Qual é a probabilidade de uma estudante portador de XX fazer uma pergunta durante a aula?

Estimativa de parâmetros:

Modelo probabilístico:

Em teoria das probabilidades e estatística, a **distribuição binomia**l é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas tais que:

Cada tentativa tem exclusivamente como resultado duas possibilidades, sucesso ou fracasso (binomial, a que se chama de tentativa de Bernoulli), e;

Cada tentativa é independente das demais, e;

A probabilidade de sucesso a cada tentativa permanece constante independente das demais, e;

A variável de interesse, ou pretendida, é o número de sucessos nas tentativas

Estimativa de parâmetros:

Qual é a probabilidade de uma estudante do sexo feminino fazer uma pergunta durante a aula?

Modelo probabilístico:

Em teoria das probabilidades e estatística, a **distribuição binomia**l é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas.

$$f(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$$

onde ***k*** é no número de eventos de sucesso (i.e., pergunta de aluna), ***n*** é o número de perguntas feitas, e ***p*** é a probabilidade deste evento ocorrer, neste caso, ele é o parâmetro (*Ø*) que queremos estimar.

Estimativa de parâmetros:

Qual é a probabilidade de uma estudante do sexo feminino fazer uma pergunta durante a aula?

Modelo probabilístico:

Em teoria das probabilidades e estatística, a **distribuição binomia**l é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas .

$$f(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$$

Observações de penguntas em PSB 2016:

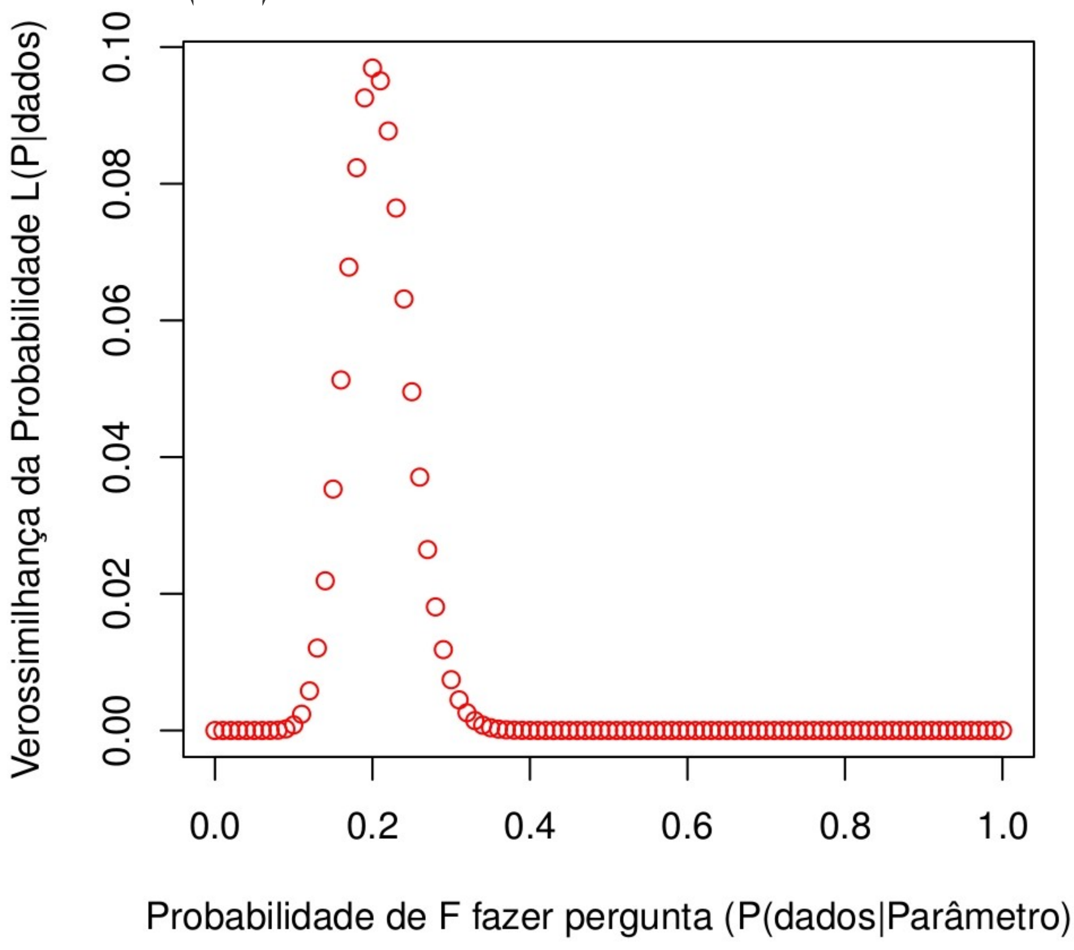
104 perguntas em 3 aulas
21 provenientes de portadores de XX

Estimativa de parâmetros:

Modelo probabilístico:

$$f(21;104,\theta)=\binom{n}{k}p^k(1-p)^{n-k}$$

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



Verossimilhanças e seleção de modelos:

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

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

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.

Fisher, R. A. 1912. On an absolute criterion for fitting frequency curves. Mess. Math. 41:155–160.

Cálculo de probabilidades:

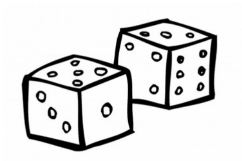
Regra 1a: A probabilidade de um evento ***A*** ocorrer é $\boldsymbol{P}_{(A)}$ = A/S, onde S é o espaço de soluções possíveis.

Regra 1b: A probabilidade de um evento ***A*** não ocorrer é $\boldsymbol{P}_{(A^{c})} = 1 - \boldsymbol{P}_{(A)}$

$\boldsymbol{P}_{(⚂)} = A/S = 1/6$

$\boldsymbol{P}_{(⚄^{C})} = \mathbf{1 - P_{(A)} = 1 - 1/6}$

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

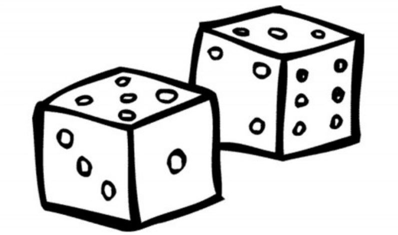


Regra 2: Se dois eventos A e B são disjuntos, então a probabilidade de qualquer um destes eventos e a soma das probabilidades dos dois eventos: $\boldsymbol{P}_{(A\ ou\ B)} = \boldsymbol{P}_{(a)} + \boldsymbol{P}_{(b)}$

Fonte: http://www.stat.yale.edu/Courses/1997-98/101/probint.htm

Probabilidades e modelos

Considere que você jogue dois dados:



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

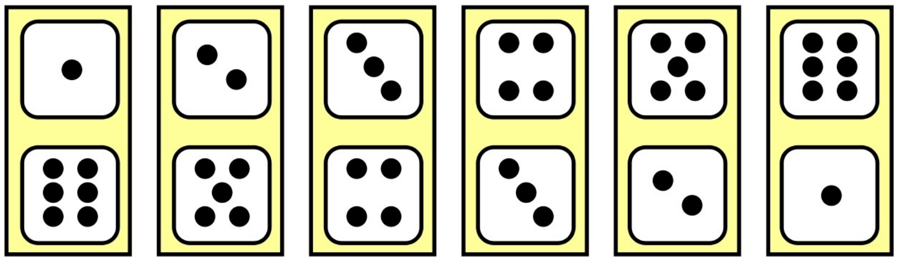
Regra 3: Se dois eventos A e B são independentes, então a probabilidade dos dois eventos ocorrerem um após o outro (i.e., intersecção de eventos, ou **probabilidade conjunta**) é o produto das probabilidades de cada evento: $\boldsymbol{P}_{(A\ e\ B)} = \boldsymbol{P}_{(A)} * \boldsymbol{P}_{(B)}$.

Combinando as regras:

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

E

1/6 * 1/6



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

OU **OU** **OU** **OU** **OU**

Fonte: Modificado de Paul O. Lewis (UCONN), 2011.

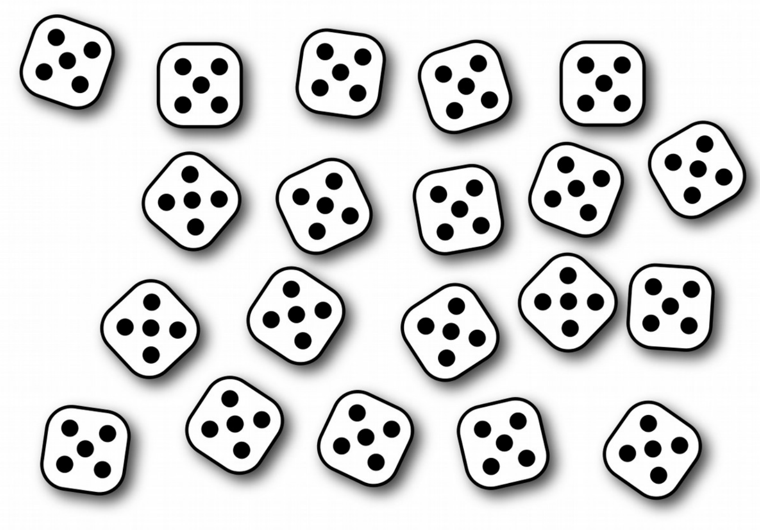
Verossimilhanças e seleção de modelos:

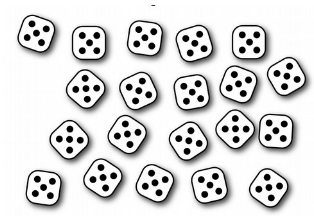
$$L(\theta|x) \approx P(x|\theta)$$

Considere:

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

$$x =$$



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

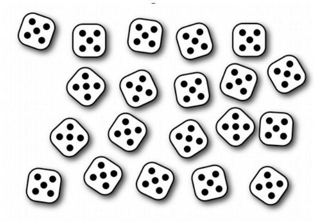
Considere que você deseja avaliar a L de dois modelos ($\theta 1$ e $\theta 2$):

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., $P_{(\text{A e B})} = P_{(\text{A})} \times P_{(\text{B})}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

Modelo 2 (θ_2) – dado duvidoso (Feito em **Brasilia/DF**) – assume que todas as faces do dado tem o número 5. Portanto, a probabilidade conjunta (i.e., $P_{(\text{A e B})} = P_{(\text{A})} \times P_{(\text{B})}$) 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.

Verossimilhanças e seleção de modelos:

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

Considere que você deseja avaliar a L de dois modelos ($\theta 1$ e $\theta 2$):

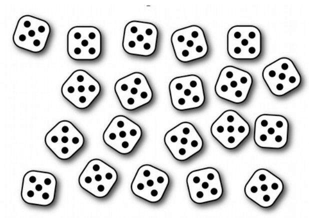
Modelo 1 (θ_1) – dado honesto – assume que cada face do dado possui um número de 1 a 6 cuja a probabilidade conjunta (i.e., $P_{(A\ e\ B)} = P_{(A)} \times P_{(B)}$) de obter 20 dados apresentando 5 é igual a $(1/6)^{20}$.

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

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

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

Verossimilhança vs. Probabilidade

$x =$  $=$ observação $= O$

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

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

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

$x =$ 

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

Verossimilhança vs. Probabilidade

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

Considere:

$\boldsymbol{O}$ = Professor ouve um barulho na goiabeira.

$\boldsymbol{H}$ = Jesus subiu na goiabeira.

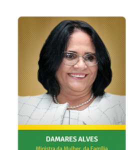
$$L(H|O) \approx P(O|H) = \text{Alta}$$

A verossimilhança desta hipótese é muito alta, pois se Jesus subuina goiabeira, 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) de que Jesus subiu na goiabeira!!!*

* **O MDH adverte**: a Ministra de Estado da Mulher, da Família e dos Direitos Humanos discorda!



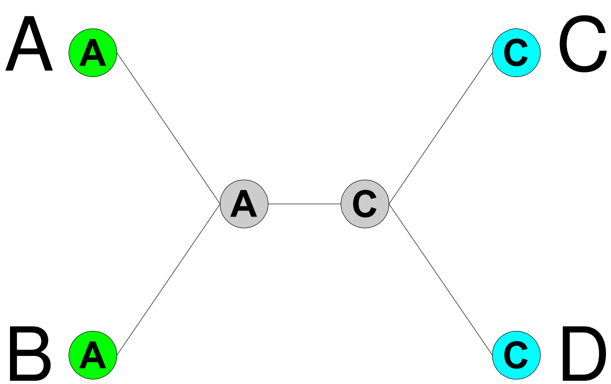
Voltemos a ideia original:

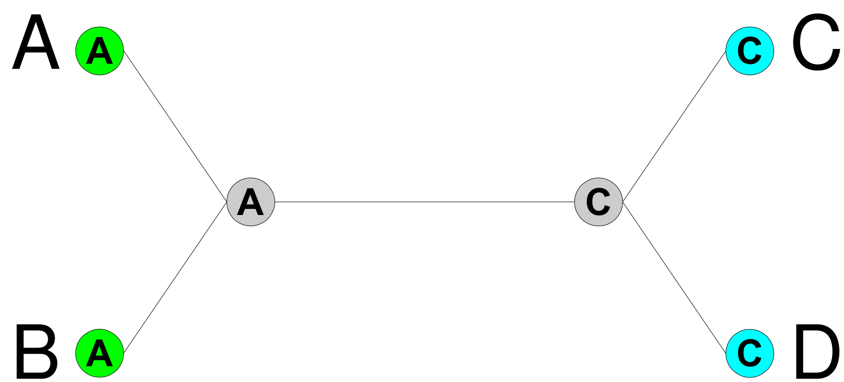
Sir Ronald Aylmer Fisher
(17/02/1890 – 29/07/1962)

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

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: TOPOLOGIA 1





Relevância da estimativa de comprimento de ramos

Voltemos a ideia original:

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

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

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:

OTU **CCACGTACGTACGTACGTACGTACGTACGTAC**

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

onde ***v*** e a “distância evolutiva entre essas sequências.

HTU **TTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:

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

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

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

Estimativa de máxima verossimilhança:

OTU **CCACGTACGTACGTACGTACGTACGTACGTAC**

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

onde ***v*** e a “distância evolutiva entre essas sequências.

HTU **TTACGTACGTACGTACGTACGTACGTACGTAC**

Voltemos a ideia original:

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

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

Estimativa de máxima verossimilhança:

probabilidade de não mudança de estado*

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

OTU CCACGTACGTACGTACGTACGTACGTACGTAC

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

HTU TTACGTACGTACGTACGTACGTACGTACGTAC

probabilidade de mudança de estado*

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

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

Voltemos a ideia original:

Estimativa de **máxima** verossimilhança:

OTU CCACGTACGTACGTACGTACGTACGTACGTAC

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

HTU TTACGTACGTACGTACGTACGTACGTACGTAC

probabilidade de não mudança de estado*

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

probabilidade de mudança de estado*

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

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

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

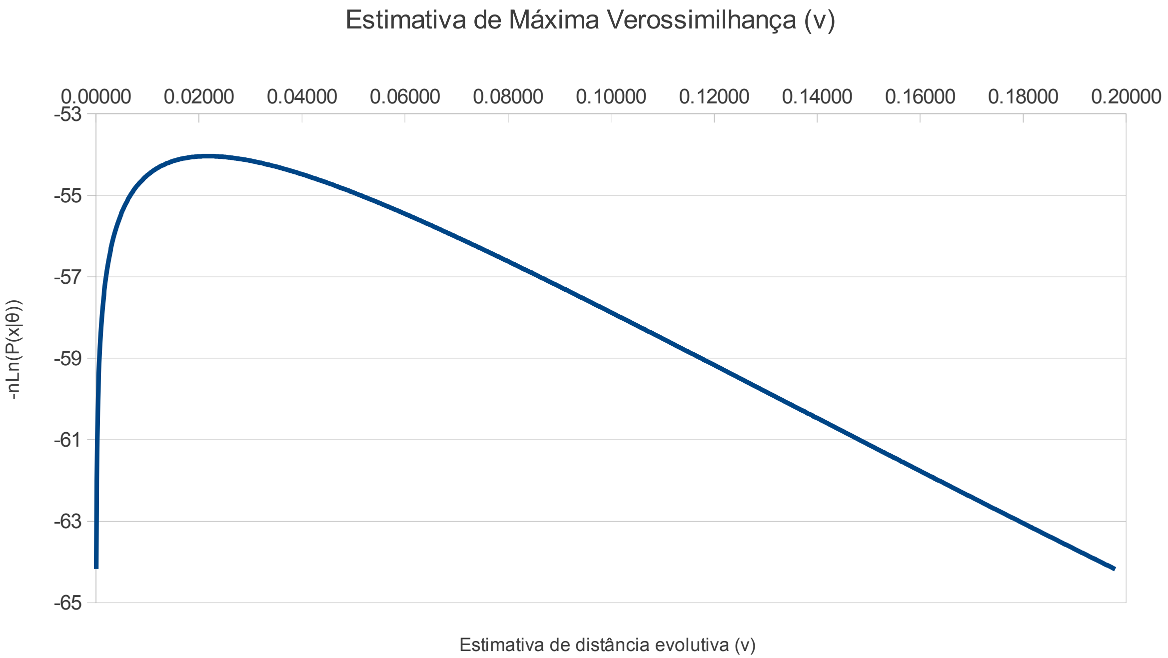
Voltemos a ideia original:

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

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

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

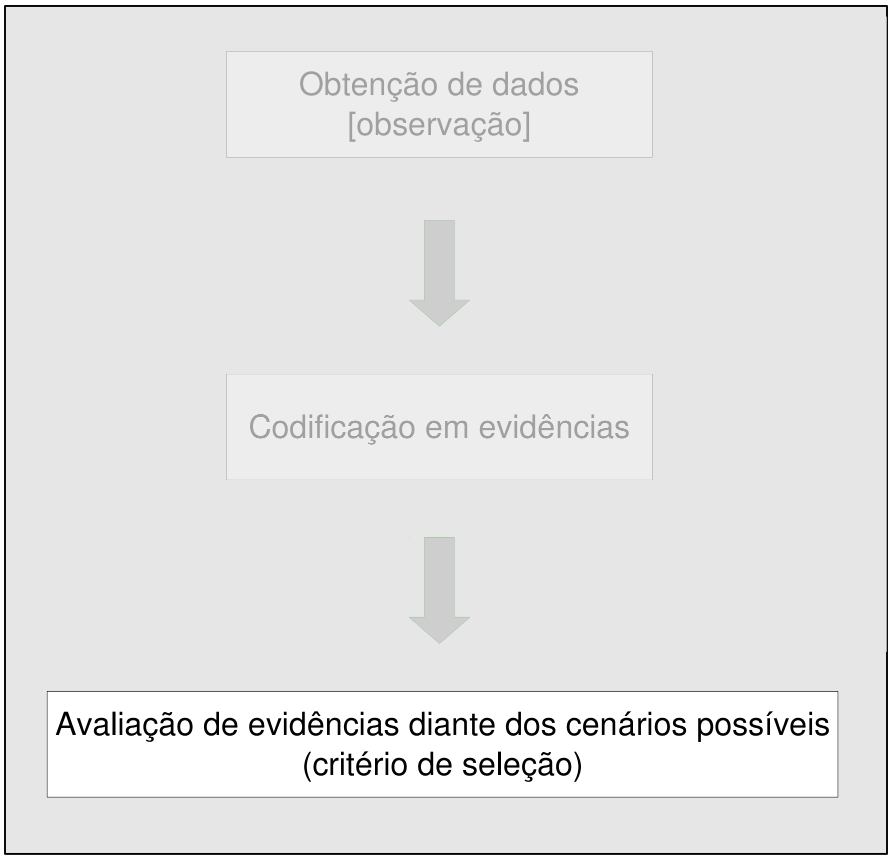
$$L=[(\frac{1}{4})*(\frac{1}{4}-\frac{1}{4}e^{-4v})]^2*[(\frac{1}{4})*(\frac{1}{4}+\frac{3}{4}e^{-4v})]^{30}$$



Valor estimado para v: 0.021752

Lógica da inferência filogenética

Avaliação e critério de seleção: soluções possíveis

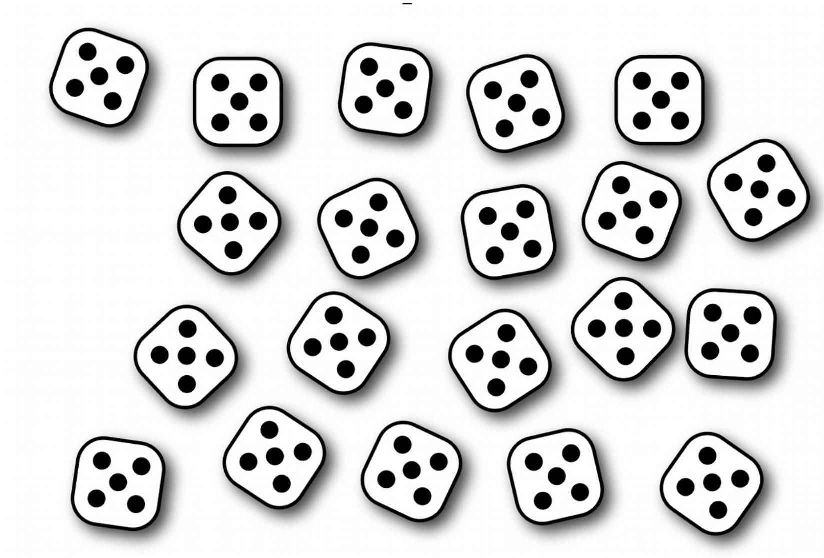


Critério de seleção: **verossimilhança (*Likelihood*)**

A probabilidade das observações calculadas utilizando **<u>um modelo</u>** nos diz o quão surpresos nós estaríamos com os dados observados

O modelo escolhido é aquele que menos surpreende!

Considere que eu jogue 20 dados sobre uma mesa e obtenha o seguinte resultados:



Fonte: Modificado de Paul O. Lewis (UCONN), 2011.

Lógica da inferência filogenética

Obtenção de dados
[observação]

Codificação em evidências

Avaliação de evidências diante dos cenários possíveis
(critério de seleção)

```
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

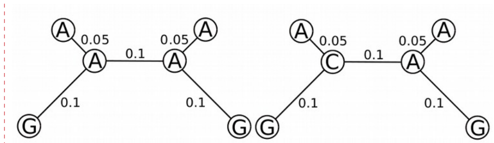
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
sp.X	1	3	2	2	1	3	0	1	2	3
sp.A	3	2	2	0	2	3	0	0	2	3
sp.B	1	1	3	0	2	1	0	0	2	3
sp.C	1	1	3	2	0	3	3	2	1	0

Probabilística (ML):

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

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

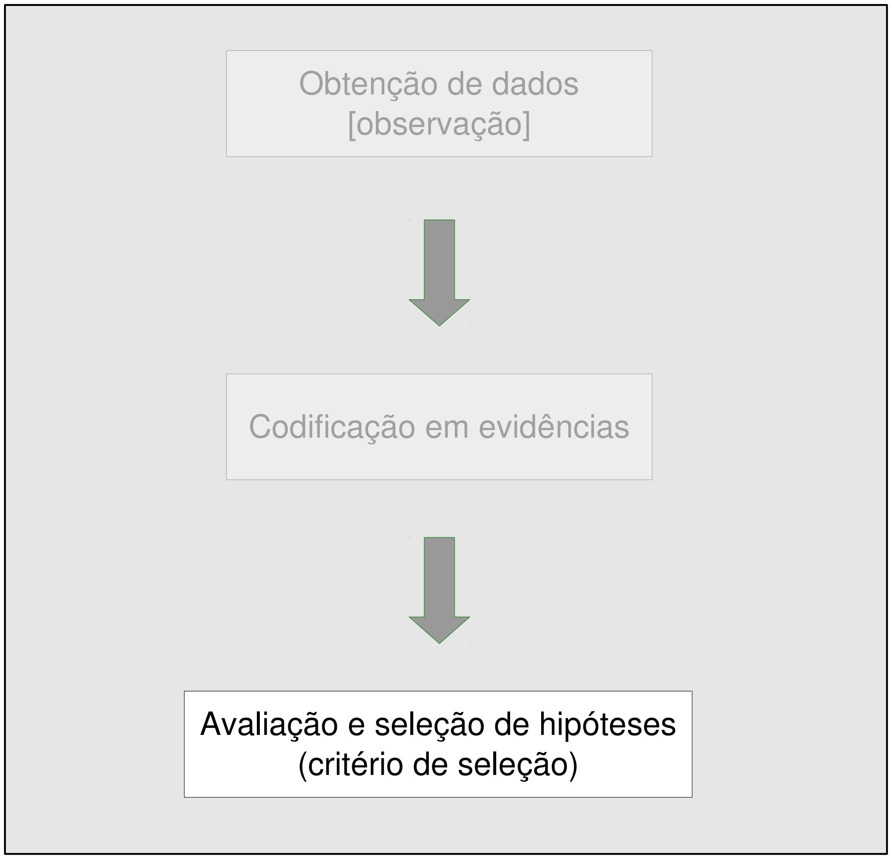
+




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

Lógica da inferência filogenética

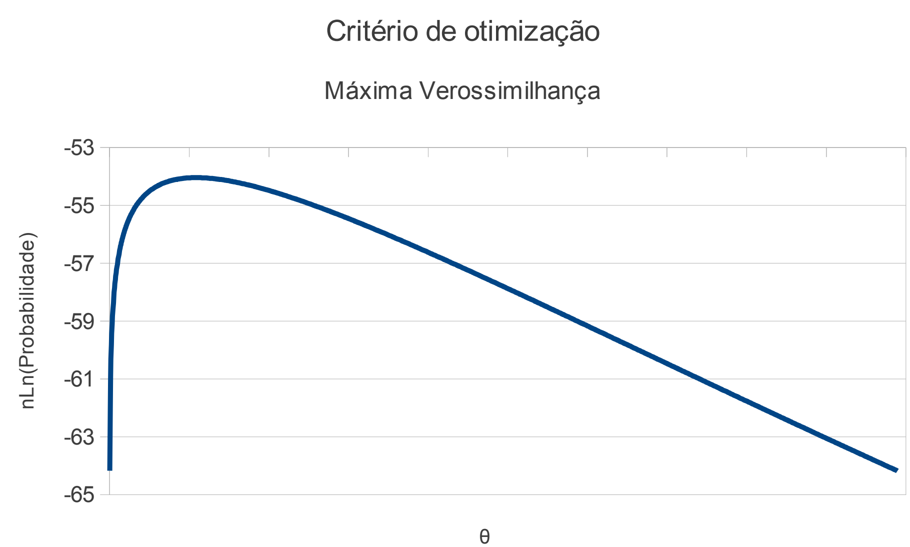
Avaliação e critério de seleção: verossimilhança máxima (ML)



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

$$L\left(\theta \middle| \begin{array}{l} \texttt{sp.X CTGGCTACGT} \\ \texttt{sp.A TGGAGTAAGT} \\ \texttt{sp.B CCTAGCAAGT} \\ \texttt{sp.C CCTGATTGCA} \end{array}\right) \approx P\left(\begin{array}{l} \texttt{sp.X CTGGCTACGT} \\ \texttt{sp.A TGGAGTAAGT} \\ \texttt{sp.B CCTAGCAAGT} \\ \texttt{sp.C CCTGATTGCA} \end{array} \middle| \theta\right)$$

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



Lógica da inferência filogenética

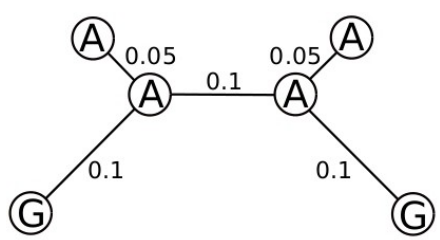
Avaliação e critério de seleção: verossimilhança máxima (ML)

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

$$L\left(\theta\,\middle|\,\begin{array}{l}\texttt{sp.X CTGGCTACGT}\\ \texttt{sp.A TGGAGTAAGT}\\ \texttt{sp.B CCTAGCAAGT}\\ \texttt{sp.C CCTGATTGCA}\end{array}\right) \approx P\left(\begin{array}{l}\texttt{sp.X CTGGCTACGT}\\ \texttt{sp.A TGGAGTAAGT}\\ \texttt{sp.B CCTAGCAAGT}\\ \texttt{sp.C CCTGATTGCA}\end{array}\,\middle|\,\theta\right)$$

Maximixação do L θ:

$$\theta =$$

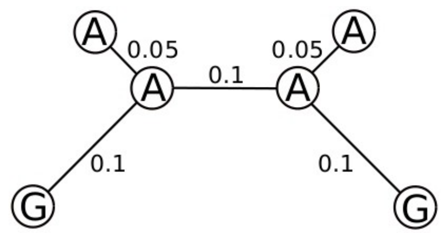


$$+ \quad GRT = \begin{bmatrix} P_{AA}(t) & P_{AC}(t) & P_{AG}(t) & P_{AT}(t) \\ P_{CA}(t) & P_{CC}(t) & P_{CG}(t) & P_{CT}(t) \\ P_{GA}(t) & P_{GC}(t) & P_{GG}(t) & P_{GT}(t) \\ P_{TA}(t) & P_{TC}(t) & P_{TG}(t) & P_{TT}(t) \end{bmatrix}$$

Lógica da inferência filogenética

Avaliação e critério de seleção: verossimilhança máxima (ML)

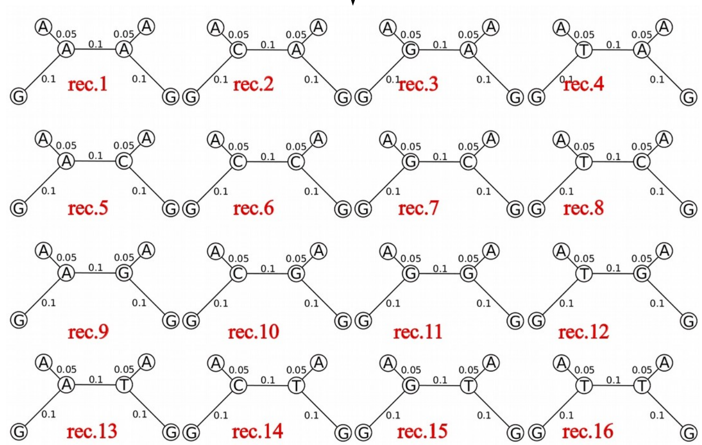
$\theta =$



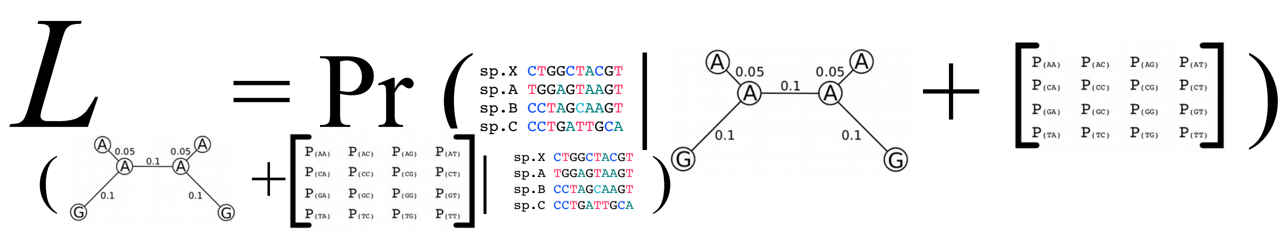
$+$

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

Possívies recontruções para HTUs



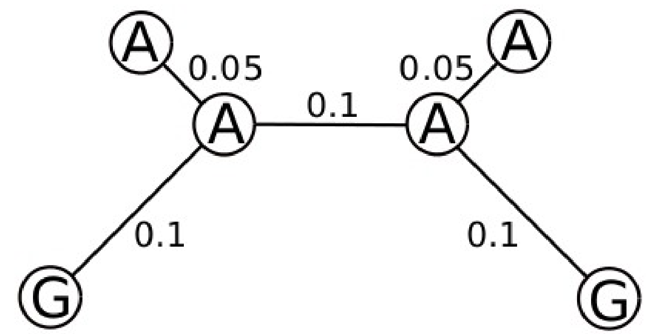
$+$ Diferentes topologias
Diferentes comprimentos de ramos



Lógica da inferência filogenética: Likelihood

Função objetiva:

$$C => P\left(\begin{matrix}\texttt{sp.X CTGGCTACGT}\\ \texttt{sp.A TGGAGTAAGT}\\ \texttt{sp.B CCTAGCAAGT}\\ \texttt{sp.C CCTGATTGCA}\end{matrix}\middle|\theta\right)$$



Modelo de substituição:

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

$$P_{ij}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\nu/3}$$

probabilidade de mudança de estado

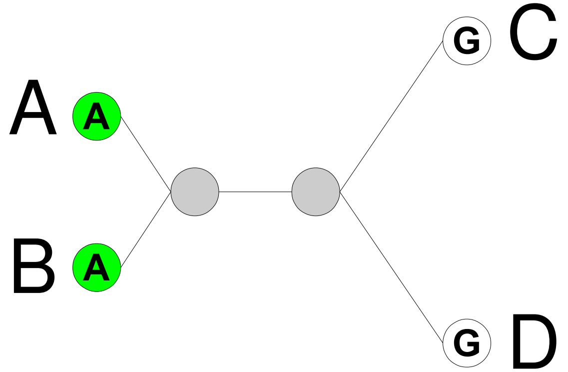
$$P_{ii}(t) = \frac{1}{4} + \frac{3}{4}e^{-4\nu/3}$$

probabilidade de não mudança de estado

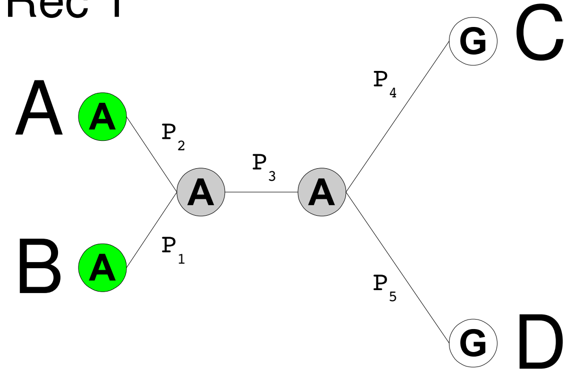
$\nu = 0.1$				
	A	**C**	**G**	**T**
A	0.9064	0.0312	0.0312	0.0312
C	0.0312	0.9064	0.0312	0.0312
G	0.0312	0.0312	0.9064	0.0312
T	0.0312	0.0312	0.0312	0.9064

$\nu = 0.05$				
	A	**C**	**G**	**T**
A	0.9516	0.0161	0.0161	0.0161
C	0.0161	0.9516	0.0161	0.0161
G	0.0161	0.0161	0.9516	0.0161
T	0.0161	0.0161	0.0161	0.9516

Cálculo de verossimilhança para Recontruções:



T1: Rec 1

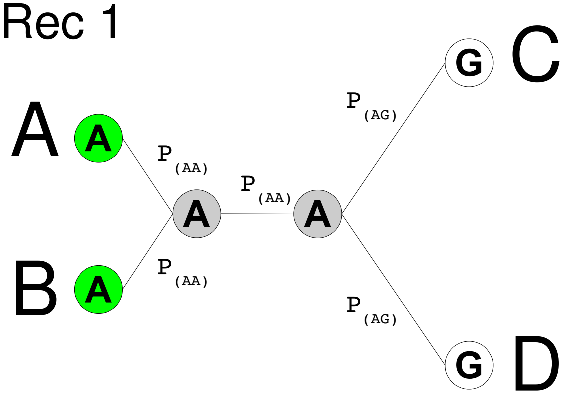


Regra 3: Iintersecção de eventos, ou probabilidade conjunta: $\boldsymbol{P_{(A\ e\ B)}} = \boldsymbol{P_{(A)}} * \boldsymbol{P_{(B)}}$.

$$Reconstrução\ 1 = P_1 * P_2 * P_3 * P_4 * P_5$$

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

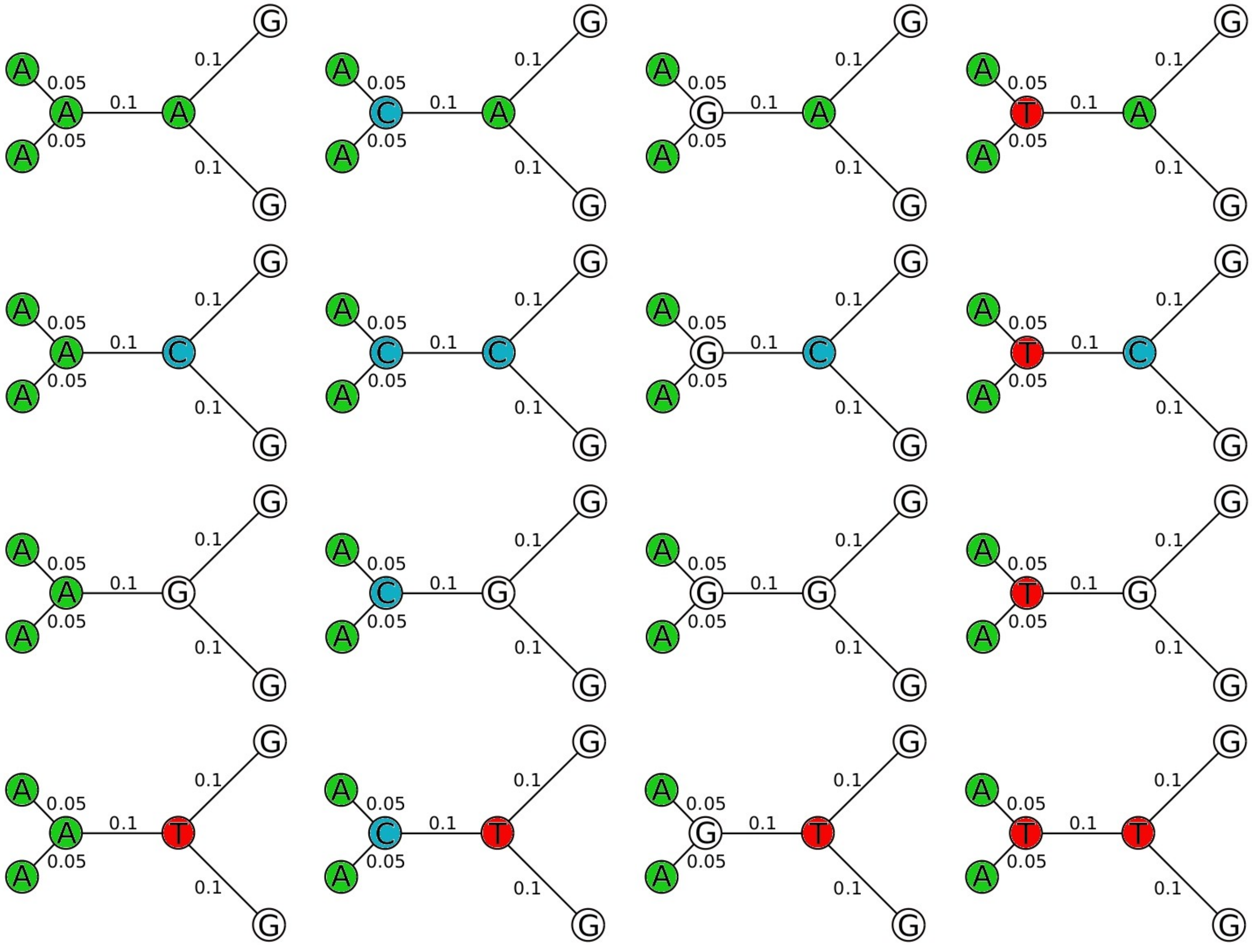
T1: Rec 1



$$P_{(T1:\ Rec\ 1)} = P_{(AA)}^{\ 3} * P_{(AG)}^{\ 2}$$

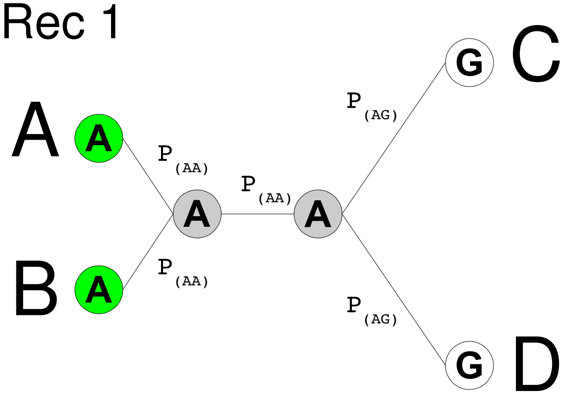
Cálculo de verossimilhança para recontruções:

Neste caso, todas as reconstruções possíveis....



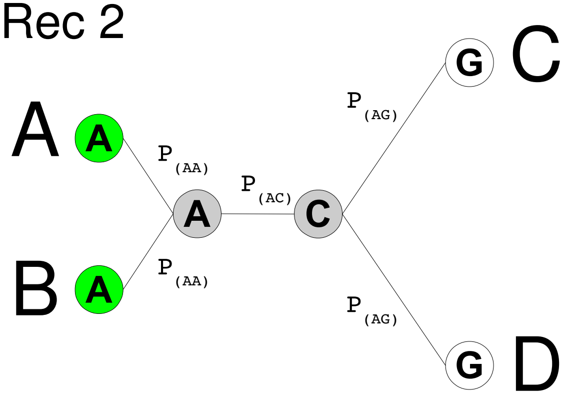
Cálculo de verossimilhança para Recontruções:

T1: Rec 1



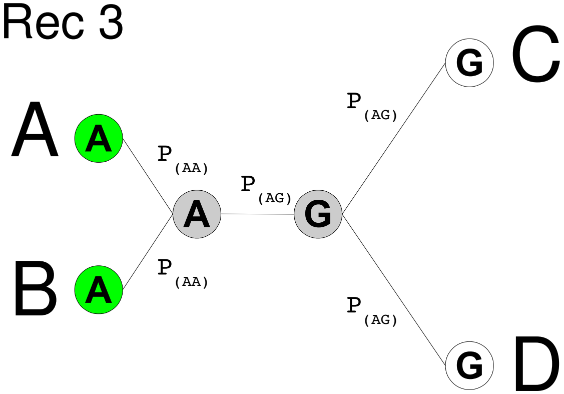
$$P_{(T1:\ Rec\ 1)} = P_{(AA)}^{3} * P_{(AG)}^{2}$$

T1: Rec 2



$$P_{(T1:\ Rec\ 2)} = P_{(AA)}^{2} * P_{(AC)} * P_{(CG)}^{2}$$

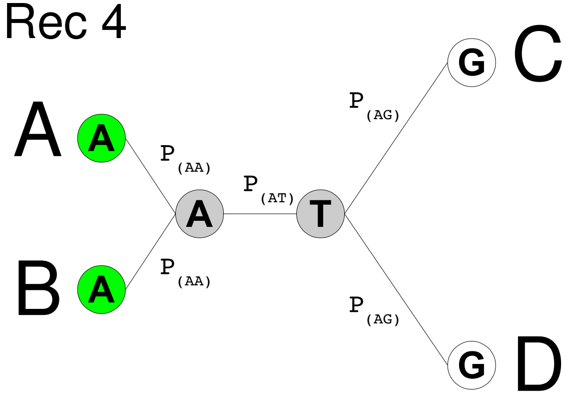
T1: Rec 3



$$P_{(T1:\ Rec\ 3)} = P_{(AA)}^{2} * P_{(AG)} * P_{(GG)}^{2}$$

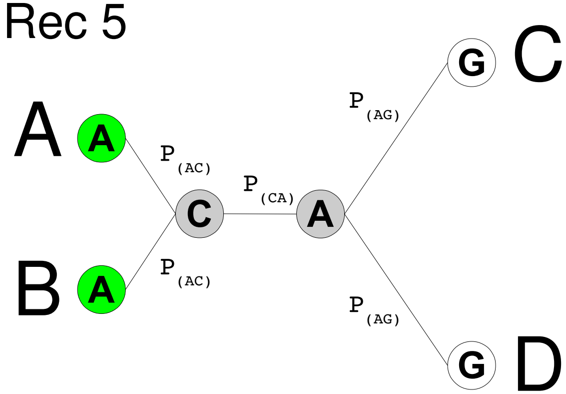
Cálculo de verossimilhança para Recontruções:

T1: Rec 4



$$P_{(\text{T1: Rec 4})} = \mathrm{P}_{(AA)}^{2} * \mathrm{P}_{(AT)} * \mathrm{P}_{(GG)}^{2}$$

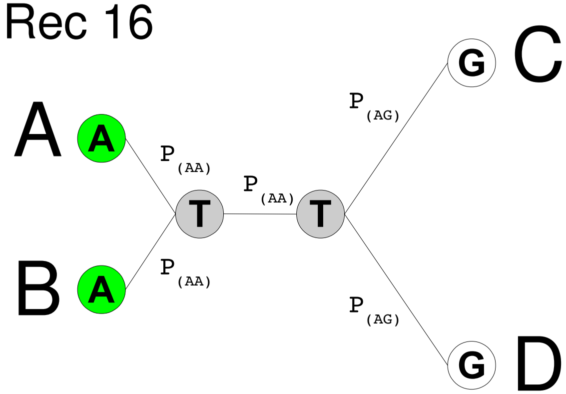
T1: Rec 5



$$P_{(\text{T1: Rec 5})} = \mathrm{P}_{(AC)}^{2} * \mathrm{P}_{(CA)} * \mathrm{P}_{(AG)}^{2}$$

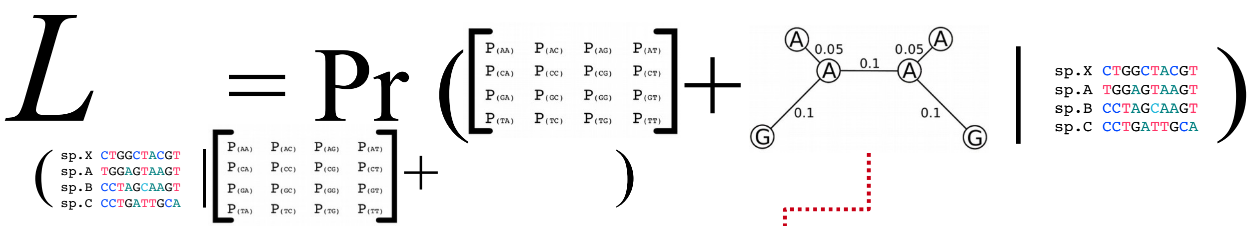
...

T1: Rec 16

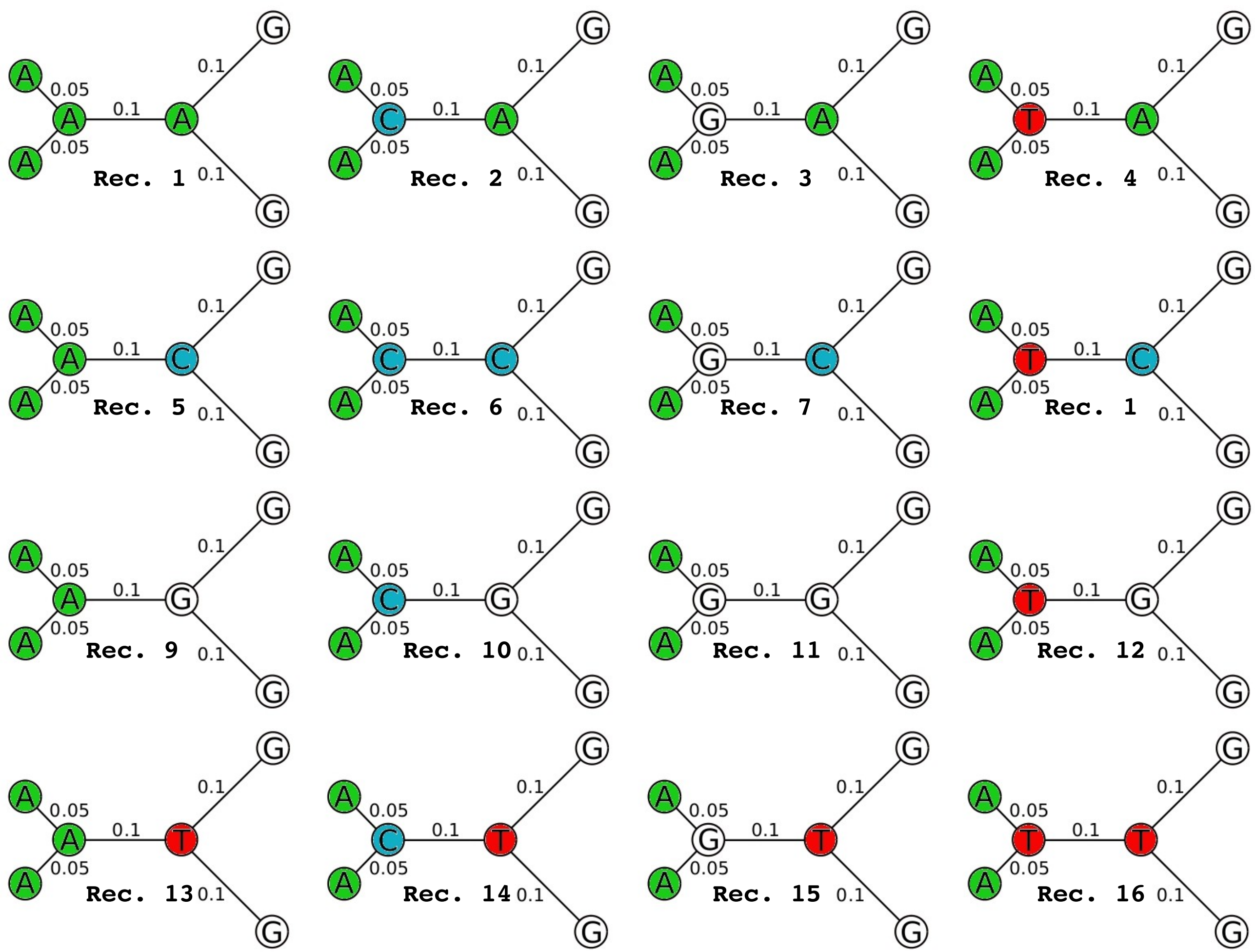


$$P_{(\text{T1: Rec 16})} = \mathrm{P}_{(AT)}^{2} * \mathrm{P}_{(TT)} * \mathrm{P}_{(TG)}^{2}$$

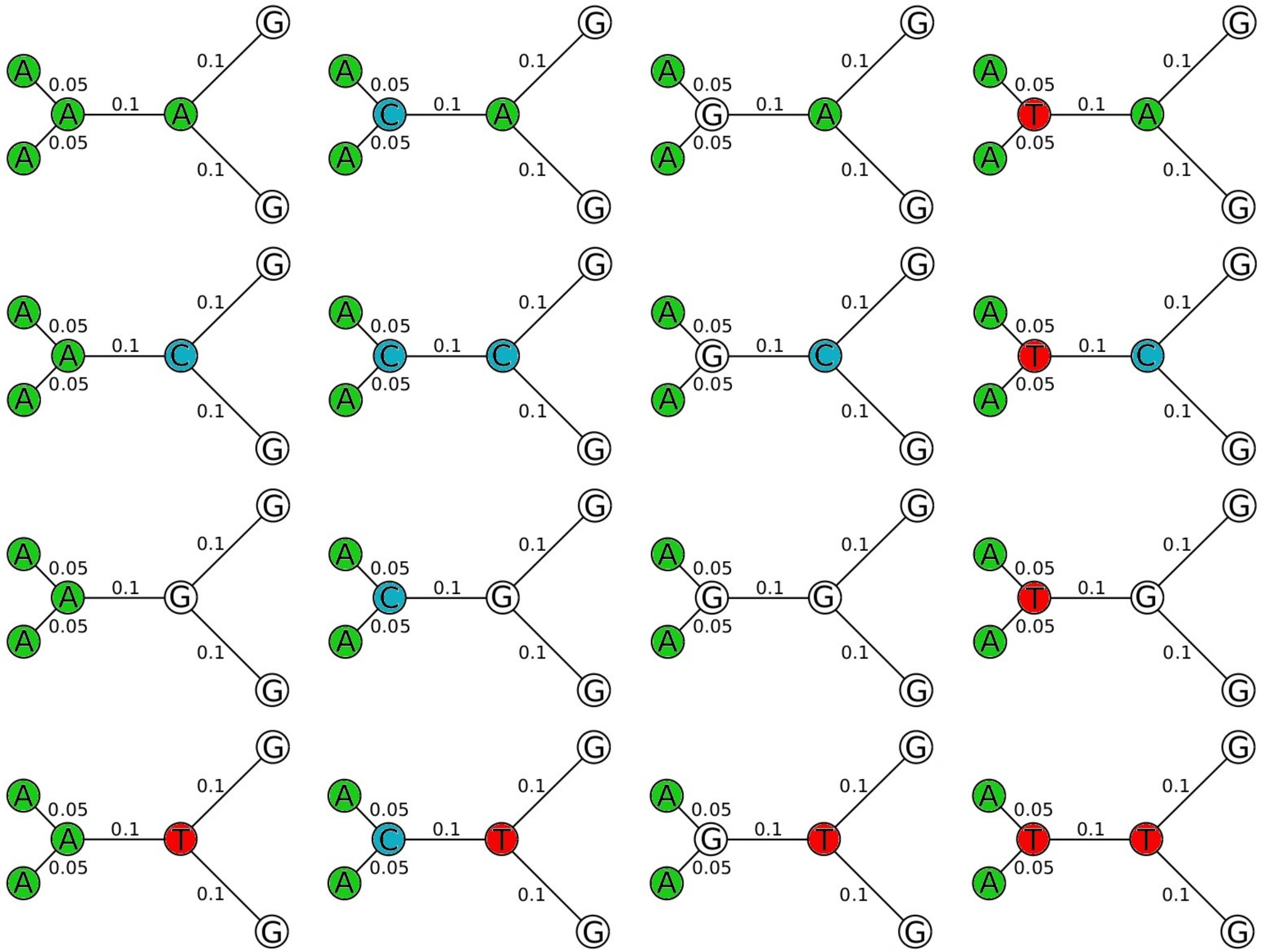
Lógica da inferência filogenética: Likelihood



Avaliação de todas as reconstruções possíveis



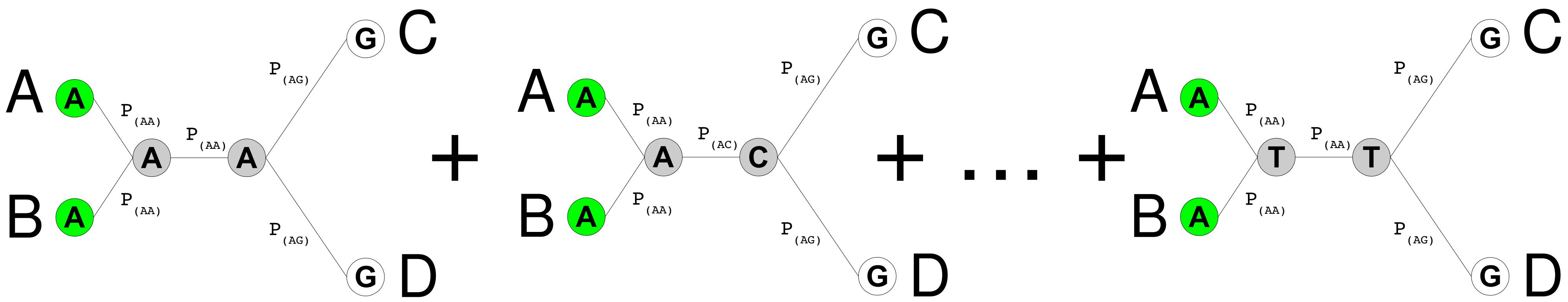
Cálculo de verossimilhança para Recontruções:



O QUE SIGNIFICA A SOMA DAS PROBABILIDADES DAS RECONSTRUÇÕES?

Cálculo de verossimilhança para topologias:

Regra 2: Eventos disjuntos: $\boldsymbol{P}_{(A\ ou\ B)} = \boldsymbol{P}_{(a)} + \boldsymbol{P}_{(b)}$



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

$$\boldsymbol{L} = \boldsymbol{L}_{(1)} * \boldsymbol{L}_{(2)} * \boldsymbol{L}_{(3)} * \ldots * \boldsymbol{L}_{(n)} = \prod_{J=1}^{N} \boldsymbol{L}_{(j)}$$

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

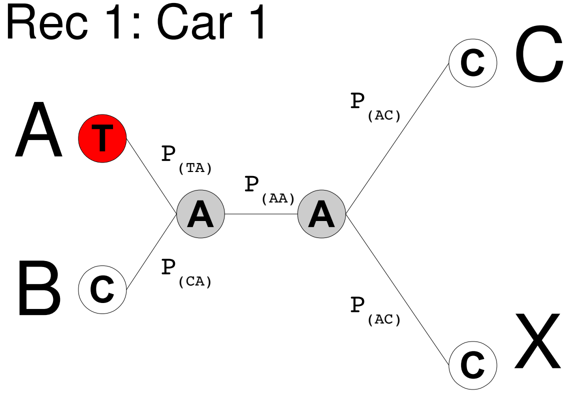
$$\ln \boldsymbol{L} = \ln \boldsymbol{L}_{(1)} + \ln \boldsymbol{L}_{(2)} + \ln \boldsymbol{L}_{(3)} + \ldots + \ln \boldsymbol{L}_{(n)} = \sum_{J=1}^{N} \ln \boldsymbol{L}_{(j)}$$

Lógica da inferência filogenética: o cálculo

Para,

```
        1 2..  j          N
sp.X    CTGGCTA...CGT
sp.A    TGGAGTA...AGT
sp.B    CCTAGCA...AGT
sp.C    CCTGATT...GCA
```

T1: Rec 1: Car 1



COMO VOCÊ FORMULARIA O CÁLCULO
PARA MAIS DE UM CARÁTER?

Lógica da inferência filogenética: modelos complexos

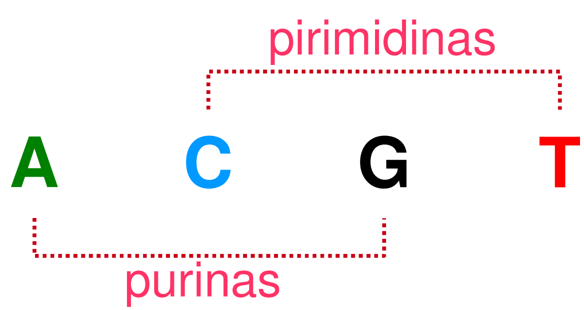
K2P

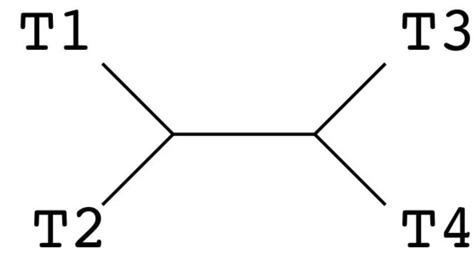
Kimura Two Parameters

Considere:

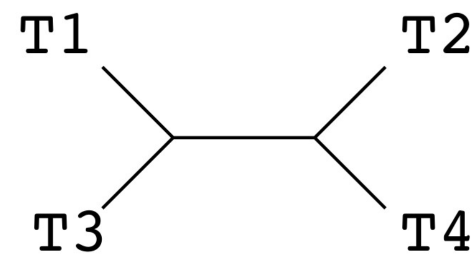
T1	**A**	**C**	**G**	**T**
T2	**C**	**G**	**T**	**A**
T3	**G**	**T**	**A**	**C**
T4	**T**	**A**	**C**	**G**

	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**	$P_{(T,G)}$	$P_{(T,T)}$

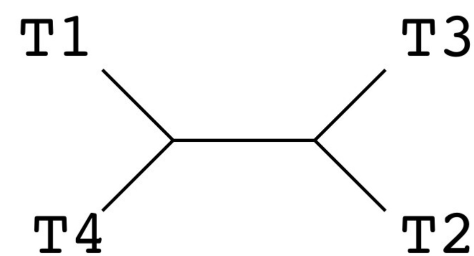




-Ln Likelihood = 22.18054
Estimated ts/tv ratio = 0.602741



-Ln Likelihood = 16.75717
Estimated ts/tv ratio = 235.857576



-Ln Likelihood = 19.52155
Estimated ts/tv ratio = 111.910550

... 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!

Escolha de critérios de otimalidade

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 otimalidade

Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

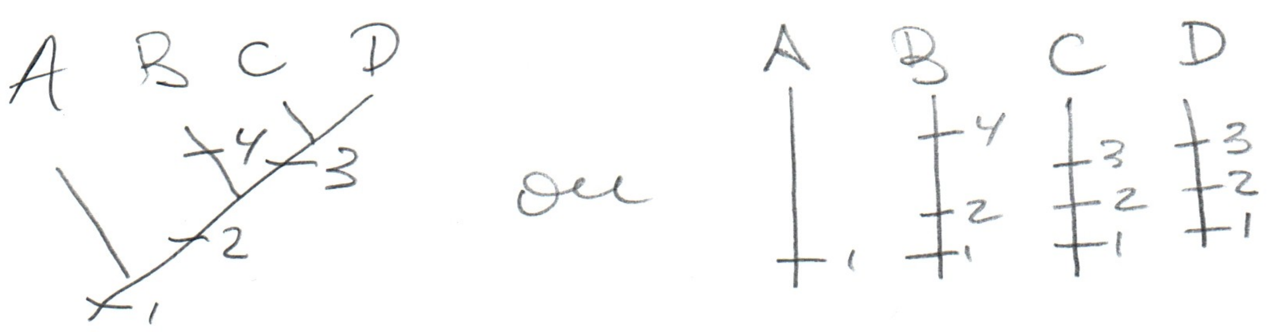
Performance

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)



Escolha de critérios de otimalidade

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

Popper --> hipóteses devem ser testadas pela busca de evidências contraditórias
> número de observação
> número de oportunidades de falsificar a hipótese
> severamente a hipótese foi testada

"Parsimonious trees have maximal explanatory power with regard to observed variation, and they do this without the accessory explanatory elements of other methods (e.g., model and tree parameters)."
(Wheeler 2012:271)

Escolha de critérios de otimalidade

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 parcimoniamente (ex. Homoplasia)

ii. Em ciência utilizada como um processo de descobrimento e teste

Iii. Pode ser defendida sem apelo ao autoritarismo a apriorismo

Escolha de critérios de otimalidade

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 otimalidade

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?

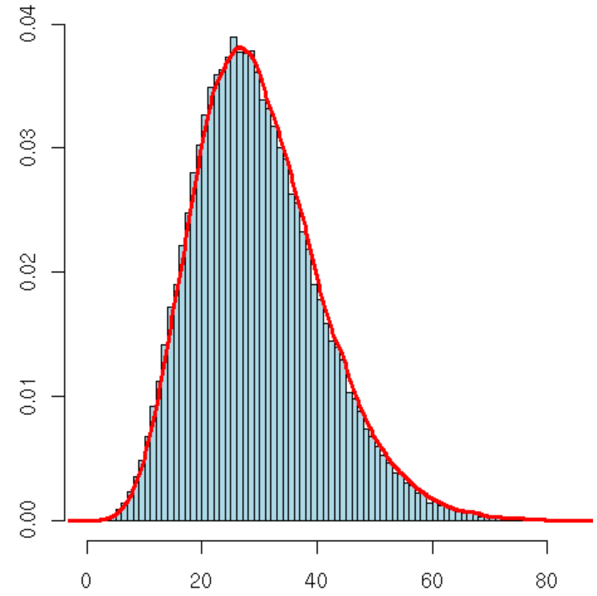
4. Teste de hipóteses e critério de otimização

Inferência hipotético-dedutiva pode ser feita via etapa de falsificação baseada em qualquer função de otimização objetiva:

Requer: hipóteses devem competir matematicamente de forma objetiva.

Questão central: a função de otimização é apropriada?

Escolha de critérios de otimalidade



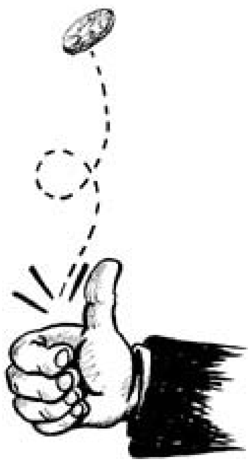
Méritos relativos dos métodos:

- **Epistemologia**
- Comportamento estatístico
- Performance
- Congruência (?)

Probabilidades: duas interpretações

i. Grau de crença:

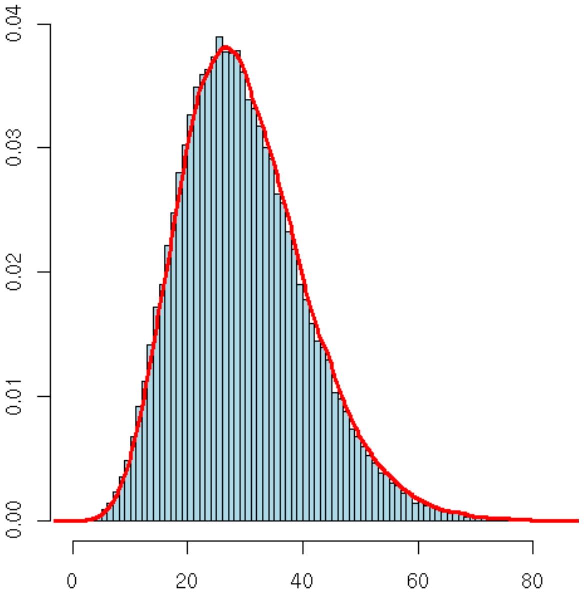
Confiança relativa em determinado resultado, considere



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

Mas a moeda é ou não é honesta!

Escolha de critérios de otimalidade



Méritos relativos dos métodos:

Epistemologia

Comportamento estatístico

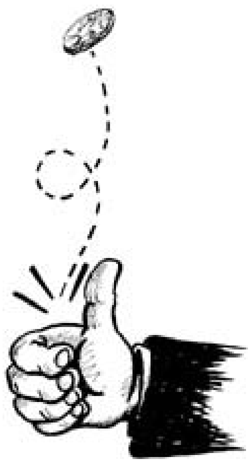
Performance

Congruência (?)

Probabilidades: duas interpretações

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

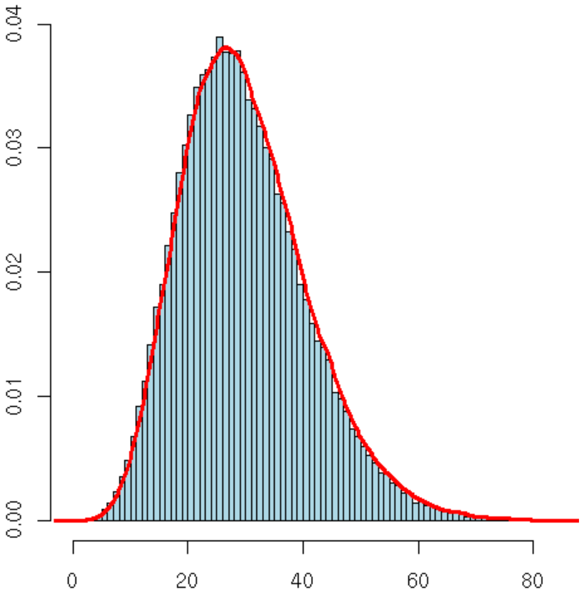
Implica em uma proposição sobre o que é ontologicamente real, considere:



"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 otimalidade



Méritos relativos dos métodos:

- **Epistemologia**
- Comportamento estatístico
- Performance
- Congruência (?)

Probabilidades:

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

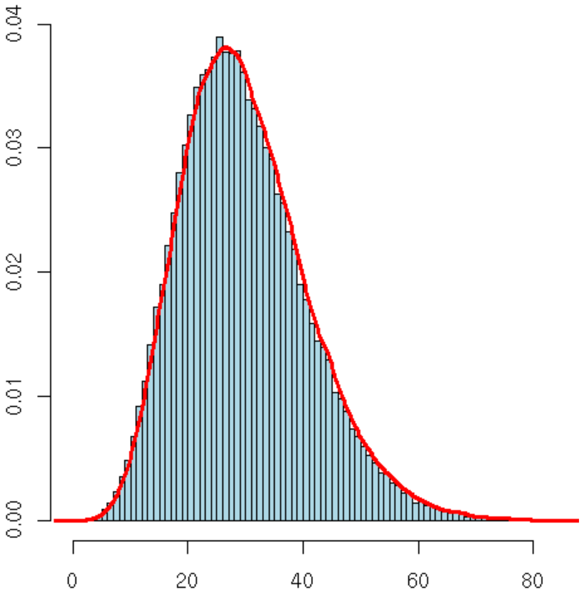
Considere:

"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 (?)



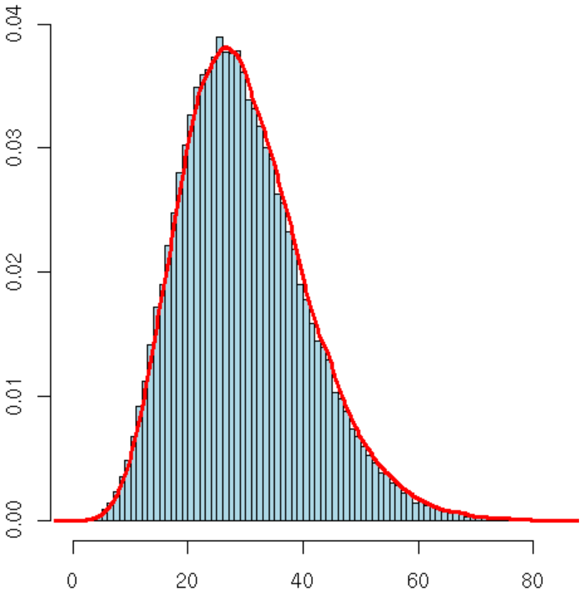
Probabilidades:

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

Considere:

"A e B são grupos-irmãos?"

Escolha de critérios de otimalidade



Méritos relativos dos métodos:

- **Epistemologia**
- Comportamento estatístico
- Performance
- Congruência (?)

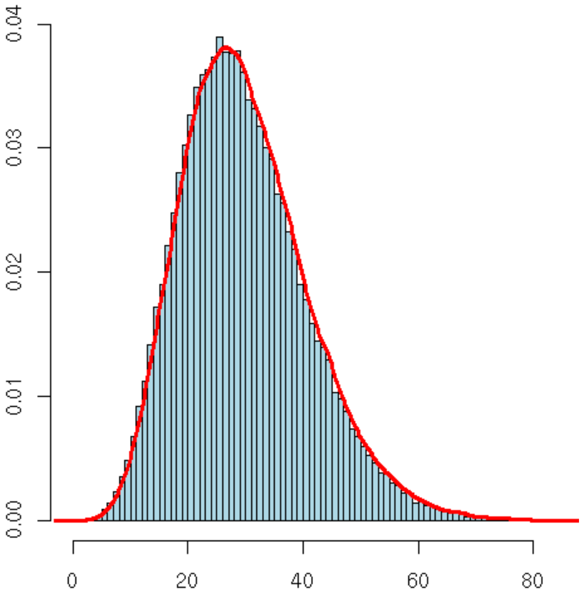
Probabilidades:

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

Considere:

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

Escolha de critérios de otimalidade



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 bilief based on empirical observation and an optimality criterion"*
>
> (Wheeler 2012:274)

Escolha de critérios de otimização

Como escolher?

Epistemologia não oferece caminho inequívoco

i. Todos os métodos podem participar do processo hipotético-dedutivo.

ii. Todos estão sujeitos à 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.