Introducción a la Inferencia Filogenética y Evolución Molecular

23-26 Junio 2008, Fac. C. Biológicas - UANL

Pablo Vinuesa (vinuesa@ccg.unam.mx)

Centro de Ciencias Genómicas-UNAM, México
http://www.ccg.unam.mx/~vinuesa/

Todo el material del curso lo puedes descargar desde:

http://www.ccg.unam.mx/~vinuesa/UANL08

- **Tema 8: Criterios de optimización II: máxima verosimilitud**

1. El criterio de máxima verosimilitud y su aplicación en filogenética
2. Estima de máxima verosimilitud de parámetros del modelo de sustitución
3. Selección de modelos de sustitución usando pruebas de razones de verosimilitud (LRTs)
4. Modelos paramétricos de evolución de secuencias de DNA – la familia GTR
5. Prácticas con PAUP*, ModelTest y proml

Métodos de reconstrucción filogenética – Máxima Verosimilitud

Máxima verosimilitud: dadas dos topologías, la que hace los datos observados más probables ("menos sorprendentes") es la preferida

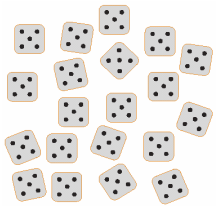
El **método de máxima verosimilitud (ML)** considera cada sitio variable del alineamiento (incluídos singletones). Bajo el criterio de ML se busca la topología que hace más verosímil el patrón de sustituciones de un alineamiento dado un modelo evolutivo explícito!

Así, para un set de datos D y una hipótesis evolutiva (topología) H, la verosimilitud de dichos datos viene dado por la expresión:

L_D=Pr(D|H) que es la probabilidad de obtener D dada H (una **probabilidad condicional**) !

Por tanto la topología que hace nuestros datos el resultado evolutivo más probable corresponde a la estima de máxima verosimilitud de la filogenia (likelihood score ó valor de verosimilitud).

- la probabilidad está relacionada con la "sorpresividad" de los datos
- Estaríamos sorprendidos de obtener este resultado, dada su bajísima probabilidad $(1/6)^{20}$ ó 1 en 3,656,158, 440,062,976!
- Pero la probabilidad depende del modelo probabilístico asumido
- En filogenética, las distintas topologías representan a los distintos modelos, y se selecciona aquel modelo que nos hace sorprendernos menos de los datos que hemos coleccionado



Máxima verosimilitud y estima de parámetros de modelos de sustitución

- La inferencia filogenética bajo el criterio de máxima verosimilitud se basa en el uso de una cantidad llamada **log-likelihood** para evaluar topologías alternativas con el fin de encontrar aquella que **maximiza este valor**.
- El **log-likelihood** es el ln de la verosimilitud, que es igual a la probabilidad de los datos observados dadas una topología particular (τ), set de longitudes de rama (υ) y modelo de sustitución (ϕ).
- Nótese que **la verosimilitud no representa la probabilidad de que un árbol sea correcto**; ésta viene determinada por la **probabilidad posterior** de la estadística bayesiana.
- Hablar de la "verosimilitud de un conjunto de datos" no es correcto ya que la verosimilitud es un función de los parámetros de un modelo estadístico, y no de los datos (*D*). **Los datos son constantes siendo el modelo lo que es variable al calcular verosimilitudes**. Se puede por lo tanto hablar de verosimilitudes como funciones de modelos o hipótesis (*H*). La verosimilitud de una hipótesis dado un set de datos es igual a la probabilidad condicional de los datos dada una hipótesis.

Formalmente: $L(H|D) = Pr(D|H) = Pr(D|\tau\upsilon\phi)$

Máxima verosimilitud y estima de parámetros de modelos de sustitución

$$L(H|D) = Pr(D|H) = Pr(D|\tau\upsilon\phi)$$

- Lo mejor es pensar en los **árboles como modelos**. La verosimilitud de una topología particular (τ) será la probabilidad de los datos dada esa topología. Cada topología tiene como parámetros las longitudes de rama (υ), y la verosimilitud de un modelo (ϕ) cambia según varíen los valores de los parámetros de longitud de rama
- Por lo tanto se puede concebir la filogenética bajo el criterio de máxima verosimilitud como un **problema de selección de modelos**. Se trata de encontar las estimas de los valores de cada parámetro del modelo y luego comparar las verosimilitudes de los distintos modelos, escogiendo el mejor (topología) en base a su verosimilitud
- La topología que hace de nuestros datos el resultado evolutivo más probable (dado un modelo de sust.) es la estima de máxima verosimilitud de nuestra filogenia. Por tanto, al contrario que bajo los criterios de optimización de MP, LS o ME, **bajo ML se trata de seleccionar modelos y parámetros que maximicen la función de optimización**.

Máxima verosimilitud y estima de parámetros de modelos de sustitución

- **Cálculo del valor de máxima verosimilitud para una sola secuencia o árbol trivial con un solo nodo**

primeros 25 nt del gen *ropB* de *Bradyrhizobium japonicum* USDA110

ATGGCGCAGCAGACATTCACCGGTC

$L = \pi_A \pi_T \pi_G \pi_G \pi_C \pi_G \pi_C \pi_A \pi_G \pi_C \pi_A \pi_G \pi_A \pi_C \pi_A \pi_T \pi_T \pi_C \pi_A \pi_C \pi_C \pi_G \pi_G \pi_T \pi_C$

$= \pi_A^{nA} \pi_C^{nC} \pi_G^{nG} \pi_T^{nT} = \pi_A^6 \pi_C^8 \pi_G^7 \pi_T^4$

$\ln L = 6 \ln(\pi_A) + 8 \ln(\pi_C) + 7 \ln(\pi_G) + 4 \ln(\pi_T)$

$\pi_A = 0.24$
$\pi_C = 0.32$
$\pi_G = 0.28$
$\pi_T = 0.16$

- A primera vista podemos sospechar que el modelo de F81 se va a ajustar mejor a los datos que el de JC69, ya que las frecuencias de nucleótidos difieren claramente de 0.25, con exceso de Cs y defecto de Ts

Máxima verosimilitud y estima de parámetros de modelos de sustitución

- **Cálculo del valor de máxima verosimilitud para una sola secuencia o árbol trivial con un solo nodo**

primeros 25 nt del gen *ropB* de *Bradyrhizobium japonicum* USDA110

ATGGCGCAGCAGACATTCACCGGTC

- Cálculo de ln*L* bajo el modelo de JC69

$\ln L = 6 \ln(\pi_A) + 8 \ln(\pi_C) + 7 \ln(\pi_G) + 4 \ln(\pi_T)$
$= 6 \ln(0.25) + 8 \ln(0.25) + 7 \ln(0.25) + 4 \ln(0.25) = -29.1$

- Cálculo de ln*L* bajo el modelo de F81

$\ln L = 6 \ln(\pi_A) + 8 \ln(\pi_C) + 7 \ln(\pi_G) + 4 \ln(\pi_T)$
$= 6 \ln(0.24) + 8 \ln(0.32) + 7 \ln(0.28) + 4 \ln(0.16) = -26.6$

$\pi_A = 0.24$
$\pi_C = 0.32$
$\pi_G = 0.28$
$\pi_T = 0.16$

-Por lo tanto el modelo de F81 se ajusta mejor a los datos (-26.6 > -29.1). Esta diferencia será tanto más notoria cuanto más larga sea la secuencia.

Máxima verosimilitud y estima de parámetros de modelos de sustitución

- **Verosimilitud del árbol más sencillo (dos nodos y una rama) bajo el modelo de JC69**

1 ——— 2

La long. de la rama equivale a la dist. evolutiva entre las hojas o nodos terminales

- Probabilidades de transición JC:
  1.- prob. de "no cambio, de i a i"
  2.- prob. de cambio, de j <-> i

$P_{ii}(\alpha t) = \Pr(\text{i en sec. 1} | \text{i en sec. 2}) = \frac{1}{4}(1 + 3e^{-4\alpha t})$
$P_{ij}(\alpha t) = \Pr(\text{j en sec. 1} | \text{i en sec. 2}) = \frac{1}{4}(1 - e^{-4\alpha t})$

GA——GG

$L = L_1 \times L_2$

$= [\Pr(G)\Pr(G\to G)][\Pr(A)\Pr(A\to G)]$
$= [\frac{1}{4}][\frac{1}{4}(1 + 3e^{-4\alpha t})][\frac{1}{4}][(\frac{1}{4} - \frac{1}{4}e^{-4\alpha t})]$
$= [1/16\,(1 + 3e^{-4\alpha t})][1/16\,(1 - e^{-4\alpha t})]$

- Cálculo de la verosimilitud por sitio (**site likelihood**) : para cada sitio L_k hay que calcular:

prob. incondic. x prob. condicional

* * * GAATCCGA ——— GGATGCGT * * *

$L = L_1 L_2 \ldots L_8 = [1/16\,(1 + 3e^{-4\alpha t})]^5\,[1/16\,(1 - e^{-4\alpha t})]^3$

$\ln L = 5 \ln[1/16\,(1 + 3e^{-4\alpha t})] + 3 \ln[1/16\,(1 - e^{-4\alpha t})]$

- Cálculo de la **verosimilitud global** para un "árbol" con 2 nodos terminales y *n* nucleótidos alineados:

$L = \prod L_k$

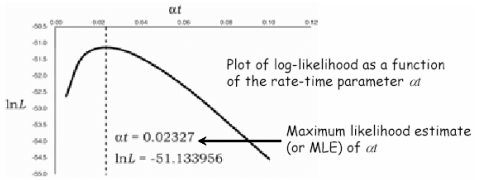
Máxima verosimilitud y estima de parámetros de modelos de sustitución

- **Estima del parámetro compuesto αt del modelo JC69 para los primeros 30 nts de la $\psi\eta$ globina de gorila y orangutan**

```
gorilla    GAAGTCCTTGAGAAATAAACTGCACACACTGG
orangutan  GGACTCCTTGAGAAATAAACTGCACACACTGG
```

$$L = \left[\left(\frac{1}{4}\right)\left(\frac{1}{4} + \frac{3}{4}e^{-4\alpha t}\right)\right]^{28}\left[\left(\frac{1}{4}\right)\left(\frac{1}{4} - \frac{1}{4}e^{-4\alpha t}\right)\right]^{2}$$

- ¿Cómo estimamos el valor de αt? La estima de máxima verosimilitud se obtiene del análisis de la función de verosimilitud, esencialmente probando diversos valores para el parámetro y determinando cual maximiza la función



$d_{JC69} = 3\alpha t$
$= 3\,(0.0237)$
$= 0.0474$

Máxima verosimilitud y estima de parámetros de modelos de sustitución

• Esquema del procedimiento del cálulo del valor de verosimilitud de un árbol con 4 OTUs

* * * GAATCCGA ——— * * * GGATGCGT

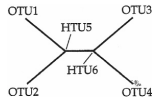
GAATCCGA
GGATGCGT

$L = L_1 L_2 \ldots L_8 =$
$= [1/16\ (1 + 3e^{-4\alpha t})]^5\ [1/16\ (1 - e^{-4\alpha t})]^3$

• En un "árbol" con sólo 2 OTUs no tenemos ningún nodo interior o ancestral. El cómputo lo realizamos directamente sobre los datos observados

• La complicación adicional que encontramos para el cálculo de verosimilitudes de árboles con > 3 OTUs radica esencialmente en que tenemos ahora nodos interiores para los que carecemos de observaciones. Se trata de unidades taxonómicas hipotéticas **HTUs**. En este caso, para calcular la verosimilitud del árbol **tenemos que considerar cada posible estado de caracter para cada nodo interior y para cada topología !!!**.

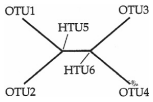
	1	2	3	4	5	6	7	8	9	...n
OTU1	A	A	G	A	C	T	T	C	A	...N
OTU2	A	G	C	C	C	T	T	C	T	...N
OTU3	A	G	A	T	A	T	C	C	A	...N
OTU4	A	G	A	G	G	T	C	C	T	...N

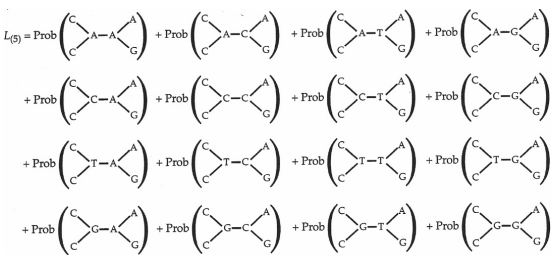


Máxima verosimilitud y estima de parámetros de modelos de sustitución

• Esquema del procedimiento del cálculo del valor de verosimilitud de un árbol con 4 OTUs

	1	2	3	4	5	6	7	8	9	...n
OTU1	A	A	G	A	C	T	T	C	A	...N
OTU2	A	G	C	C	C	T	T	C	T	...N
OTU3	A	G	A	T	A	T	C	C	A	...N
OTU4	A	G	A	G	G	T	C	C	T	...N





• Para 4 OTUs existen 3 topologías posibles.

Por ello hemos de repetir este cálculo para cada una de ellas con el fin de encontrar la topol. más verosímil

$$L = L_{(1)} \times L_{(2)} \times L_{(3)} \times \ldots \times L_{(n)} = \prod_{i=1}^{n} L_{(i)}$$

$$\ln L = \ln L_{(1)} + \ln L_{(2)} + \ln L_{(3)} + \ldots + L_{(n)} = \sum_{i=1}^{n} \ln L_{(i)}$$

• La **verosimilitud para cada sito** representa la suma sobre todas las posibles asignaciones de estados de caracter en todas las ramas interiores de un árbol. La **verosimilitud total** es el producto de las veros. por sitio.

Máxima verosimilitud y estima de parámetros de modelos de sustitución

• La inferencia filogenética bajo el criterio de máxima verosimilitud implica
MUCHISIMO TRABAJO COMPUTACIONAL
(=> mucho tiempo de trabajo de procesador)

• Las verosimilitudes globales han de ser maximizadas para cada topol. Para ello necesitamos:
- encontrar EMV para cada long. de rama y cada parámetro del modelo de sust.
- ello implica calcular la verosimilitud global muchas, pero que muchas veces

• En la práctica los **árboles de ML se estiman en múltiples ciclos**, en los que se van **optimizando secuencialmente los diversos parámetros** del modelo de sustitución y longitudes de rama. La estima conjunta de todos los parámetros se hace computacionalmente prohibitiva

• Por lo general **se comienzan estos ciclos partiendo de una topología** obtenida por un método rápido, tal como **NJ o MP**. Sobre esta topología se ajustan los valores de los parámetros del modelo. A continuación se emplea algún método de reajuste de topología (branch swapping) y se ajustan las longitudes de rama, cerrando un ciclo. En múltiples ciclos consecutivos se va optimizando la topología y long. de rama, **hasta que convergen en la estima de máxima verosimilitud global**

Máxima verosimilitud y estima de parámetros de modelos de sustitución

1. La relevancia e impacto de los modelos de evolución en filogenética y evol. molecular

- Los modelos no son sólo importantes por sus consecuencias en la estima filogenética, sino que además lo son porque la **caracterización del proceso evolutivo** a nivel molecular es **objeto de estudio en sí mismo**, en el ámbito de la evolución molecular.

- Los modelos de evolución **son herramientas poderosas** siempre y **cuando**, a pesar de las simplificaciones que hacen, **describan adecuadamente las características más salientes de los datos y permitan hacer predicciones precisas** sobre el problema bajo estudio.

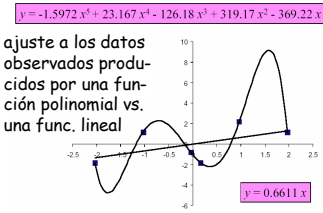
- **El rendimiento de un método se maximiza cuando se satisfacen los supuestos subyacentes.**

- Es conveniente por tanto seleccionar el modelo más adecuado para cada set particular de datos y cuantificar el ajuste de los datos al modelo seleccionado.

Máxima verosimilitud y estima de parámetros de modelos de sustitución

2. Selección de modelos de sustitución de secuencias de DNA

- En términos generales modelos complejos se ajustan a los datos mejor que los simples. Idealmente **se ha de seleccionar un modelo los suficientemente complejo** (rico en parámetros) como **para describir adecuadamente las características más notables del patrón de sust.** del set de datos, pero no sobreparametrizado para **evitar colinearidad de parámetros (redundancia),** tiempos excesivamente largos de cómputo y estimas poco precisas de los parámetros por excesiva varianza.



ajuste a los datos observados producidos por una función polinomial vs. una func. lineal

- **añadir parámetros** a un modelo generalmente mejora su ajuste a los datos observados
- **modelos infra-parametrizados** conducen a un pobre ajuste a los datos observados
- **modelos supra-parametrizados** conducen a una pobre predicción de eventos futuros
- existen métodos estadísticos para **seleccionar modelos ajustados** a cada set de datos

Máxima verosimilitud y estima de parámetros de modelos de sustitución

2. Selección de modelos de sustitución de secuencias de DNA y Proteína

- Se deben de usar **pruebas estadísticas para seleccionar el modelo que mejor se ajusta a los datos de entre los disponibles.** Este ajuste de los modelos a los datos puede ser evaluado usando pruebas de razones de verosimilitud (likelihood ratio tests, **LRTs**) o usando criterios de información de Akaike o bayesianao (**AIC** y **BIC**, respectivamente). Se puede usar una prueba de LRT para evaluar la capacidad que tiene un modelo particular en ajustar los datos.

- Idealmente debemos de seleccionar el mejor modelo para cada gen o región genómica que queramos analizar. No conviene hacerlo para una supermatriz de alineamientos concatenados. El uso de **modelos particionados** en los que se ajusta el modelo para cada posición de los codones, por cada gen a analizar, resultan generalmente en ajustes globales significativamente mejores que **modelos promediados** para cada gen.

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Prueba de razón de verosimilitudes (LRT)

- Una manera natural y muy usada de comparar el ajuste relativo de dos modelos alternativos a una matriz de datos es contrastar las verosimilitudes resultantes mediante la prueba de razones de verosimilitud (RV) ó likelihood ratio test (LRT):

$$\Delta = 2(\log_e L_1 - \log_e L_0)$$

donde L_1 es el valor de ML global para la hipótesis alternativa (modelo más rico en parámetros) y L_0 es el valor de ML global para la hipótesis nula (el modelo más simple).
$\Delta >= 0$ **siempre**, ya que los parámetros adicionales van a dar una mejor explicación de la variación estocástica en los datos que el modelo más sencillo.

- Cuando los modelos a comparar están anidados (L_0 es un caso especial de L_1) el estadístico Δ sigue aproximadamente una **distribución χ^2 con q grados de libertad, donde q = diferencia entre el no. de parámetros libres entre L_1 y L_0 .**

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Prueba de razón de verosimilitudes (LRT)

- El LRT es por tanto una prueba estadística para cuantificar la bondad relativa de ajuste entre dos modelos anidados. Veamos un ejemplo. Vamos seleccionar entre los modelos JC69, F81, HKY85 y TrN93 para el set de datos de mtDNA-primates.nex, considerando sólo las regiones codificadoras y eliminando Lemur_catta, Tarsius_syrichta y Saimiri_scireus y usando un árbol NJ sobre el cual estimar parámetros

Modelo	-lnL
JC69	3585.54820
F81	3508.04085
HKY85	3233.34395
TrN93	3232.29439

- ¿ Qué podemos concluír de estos valores de -ln$L$ en cuanto a la **importancia relativa de los parámetros** considerados por estos modelos en cuanto al nivel de ajuste a los datos que alcanzan ?

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Prueba de razón de verosimilitudes (LRT)

Modelo	-lnL	H_0 a rechazar (o hipótesis anidadas a evaluar)
JC69	3585.54820	
F81	3508.04085	1. igual frec. de bases
HKY85	3233.34395	2. Ti = Tv
TrN93	3232.29439	3. tasas de Ti iguales
		...

modelos	diff. GL = *q*	X^2	*P*	
JC-F81	3 - 0 = 3	155	0	
JC-HKY85	4 - 0 = 4	704.4	0	
JC-TrN	5 - 0 = 5	706.4	0	
F81-HKY85	4 - 3 = 1	549.4	0	← Por lo tanto el modelo seleccionado es el **HKY**
F81-TrN	5 - 3 = 2	551.4	0	
KHY-TrN	5 - 4 = 1	2.1	0.15	←

http://www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html

Máxima verosimilitud y estima de parámetros de modelos de sustitución

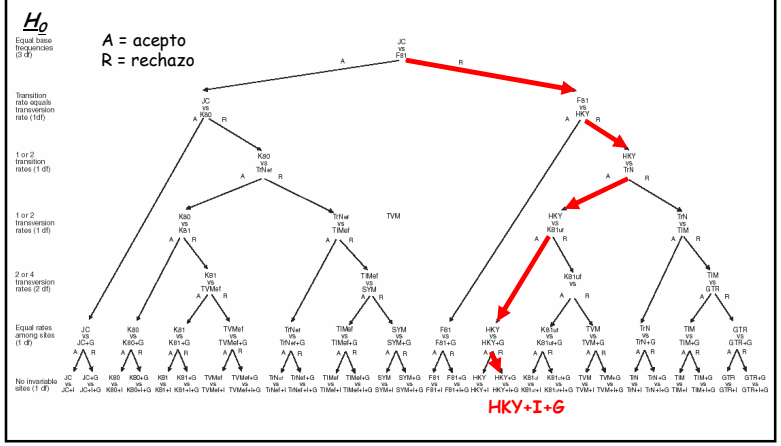
3. Prueba de razón de verosimilitudes (LRT)

Modelo	-lnL	H_0 a rechazar (o hipótesis anidadas a evaluar)
HKY85	3233.34395	
HKY85 +*G*	3145.29031	1. tasa homogénea de sust entre sitios
HKY85 +I+*G*	3142.36439	2. no existe proporción de sitios invariantes

modelos	diff. GL = *q*	X^2	*P*	
HKY85-vs. +*G*	1	176	0	← Por lo tanto el modelo seleccionado es el **HKY+*G*** si tomamos 0.01 como punto de corte, o **HKY+I+*G*** si usamos alfa = 0.05.
HKY85+*G* vs. I+*G*	1	5.85	0.015	←

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Esquema jerárquico de efectuar LRTs partiendo desde el modelo más sencillo (JC69)



Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Resumen de algunos modelos y sus parámetros libres

- Dado que en los modelos de sust. de DNA la tasa promedio de sustitución se considera = 1 y los parámetros de tasa relativa se escalan de tal manera que la tasa promedio de sust. en equilibrio = 1, el modelo más sencillo (JC69) no tiene ningún parámetro libre, dado que el único parámetro (α) a estimar valdrá $\frac{1}{4}$ en este contexto.

Modelo	características	no. de parámetros libres
JC	nst= 1 basefreq= equal	0
F81	nst=1 basefreq=uneq	3 para las *frec. de bases*
K2P	nst=2 basefreq=eq	1 para el *tratio* (ti/tv)
HKY85	nst=2 basefreq=uneq	4 (1 para *tratio* y 3 para frec. de bases)
TrN93	nst=3 basefreq=uneq	5 (2 tasas de ti y 3 para freq de bases)
GTR	nst=6 basefreq=uneq	8 (5 para tasas de subst y 3 para freq. de bases)

proporción de sitios invariantes (I)	1 parámetro libre adicional para pinv
distribución gamma (*G*)	1 parámetro libre adicional para *G*
ambos combinados (I+*G*)	2 parámetros libres adicionales

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Selección de modelos usando criterios de información

- LRT compara pares de modelos anidados. Los criterios de información como el **Akaike information criterion** (**AIC**) y **Bayesian information criterion** (**BIC**) **comparan simultáneamente todos los modelos en competición** y **permiten seleccionar modelos aunque no sean anidados.**

- Se trata nuevamente de incorporar tanta complejidad (parámetros) al modelo como requieran los datos. La verosimilitud para cada modelo es penalizada en función del número de parámetros: **a mayor cantidad de parámetros mayor penalización.**

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Selección de modelos usando criterios de información

- **AIC.** Es un estimador no sesgado del parámetro de contenido de información de Kullback-Leibler, el cual es una **medida de la información perdida al usar un modelo para aproximar la realidad.** Por tanto, **a menor valor de AIC mejor ajuste** del modelo a los datos. Al penalizar por cada parámetro adicional, **considera tanto la bondad de ajuste como la varianza asociada a la estima de los parámetros.**

$AIC_i = -2\ln L_i + 2 N_i$

N_i = no. de parámetros libres en el modelo i

L_i = verosimilitud bajo el modelo i

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Selección de modelos usando criterios de información: AIC

- Se pueden usar los **estadísticos de diferencias en AIC** (Δ_i) y **ponderaciones de Akaike para cuantificar el nivel de incertidumbre en la selección del modelo.** Las Δ_i son AICs re-escalados con respecto modelo con el AIC más bajo (minAIC), de modo que

$\Delta_i = AIC_i - \text{minAIC}$.

Las Δ_i son fáciles de interpretar y permiten ordenar los los modelos cadidatos. Así, **modelos con Δ_i en un rango de 1-2 con respecto al modelo ganador tienen un soporte sustancial** y deben de ser considerados como modelos alternativos. Modelos con Δ_i **en un rango de 3-7** con respecto al modelo ganador tienen un **soporte significativamente inferior**, y modelos con Δ_i **> 10 carecen de soporte.**

Máxima verosimilitud y estima de parámetros de modelos de sustitución

3. Selección de modelos usando criterios de información: AIC

- Las **ponderaciones o pesos de Akaike (w_i) son los AIC relativos normalizados para cada modelo en competición y pueden ser interpretados como la probabilidad de que un modelo es la mejor abstracción de la realidad dados los datos**. Para R modelos candidatos a evaluar:

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{r=1}^{R}\exp\left(-\frac{1}{2}\Delta_r\right)}$$

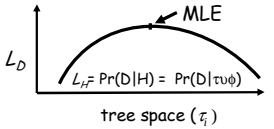
- Una aplicación muy útil de los w_i es que la inferencia se puede promediar a partir de los modelos que muestran valores de no w_i triviales. Así, una estima del valor del parámetro a de la distribución gamma promediada a partir de varios modelos se calcularía así:

$$\hat{\bar{\alpha}} = \sum_{i=1}^{R} w_i \hat{\alpha}_i$$

También podemos **reconstruir filogenias bajo los distintos modelos con peso significativo y combinar los árboles resultantes acorde a sus pesos de Akaike**. Esta estrategia es particularmente útil en un contexto bayesiano.

Criterios de optimización: la alteranativa Bayesiana

- **Aproximaciones tradicionales (matrices de distancia, ME, ML, MP)**
  - la búsqueda tiene por objetivo encontrar la topología óptima (**estima puntual**)
  - no pueden establecer el soporte relativo de las biparticiones a partir de una única búsqueda



- Aproximación Bayesiana
  - no busca una solo topología óptima sino una **población de árboles muestreados en función de su probabilidad posterior** (algoritmos MCMC)
  - la muestra de árboles obtenidos en una sola sesión de "búsqueda" es usada para valorar el soporte de cada split en términos de propabilidad posterior