Pinvar in PAUP*

Specify either a particular value for pinvar:

\[
\text{lset pinvar}=0.3;
\]

Tell PAUP* to estimate it:

\[
\text{lset pinvar}=\text{estimate};
\]

Or, if you've already estimated it, tell PAUP* to use the value it just calculated:

\[
\text{lset pinvar}=\text{previous};
\]
Discrete Gamma (G) model

Mixture model:

\[ L_i = \left( \frac{1}{3} \right) P_r(D_i | r_1) + \left( \frac{1}{3} \right) P_r(D_i | r_2) + \left( \frac{1}{3} \right) P_r(D_i | r_3) \]

\[ P_r(D_i | r_j) = \frac{1}{\Gamma(\alpha)} \left( \frac{b}{\alpha} \right)^{\alpha} \left( \frac{1}{2} \right)^{\beta} \]

\[ \Gamma(a, b) \]

\[ \text{mean} = \frac{ab}{\alpha} = 1.0 \]

\[ b = \frac{1}{a} \]


Gamma distributions

\[ \text{mean} = 1 \]

axis labels

\[ \text{shape} = \frac{1}{\text{variance}} \]

\[ \alpha = 1 \quad \text{(very little heterogeneity)} \]

\[ \alpha = 1 \]

\[ \alpha = 0.1 \]

\[ \alpha (\frac{1}{\alpha})^2 = \frac{1}{\alpha} \]

\[ \text{Gamma}(\alpha, b) \]

\[ 2b = \text{mean} \]

\[ ab^2 = \text{variance} \]
Relative rates in 4-category case

\[ r_1, r_2, r_3, r_4 \]

\[ \approx 1.0 \]

\[ \text{shape} = 1 \]

\[ ncat = 4; \]

\[ \gamma \text{plot} \]

\[ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{3}{4} \]

\[ 0, 0.5, 1, 1.5, 2, 2.5 \]

\[ 0.137, 0.477, 1.000, 2.386 \]
Discrete gamma rate heterogeneity in PAUP* 

To use gamma distributed rates with 4 categories:

\[ \text{lset rates=gamma ncat=4;} \]

To estimate the shape parameter:

\[ \text{lset shape=estimate;} \]

To combine pinvar with gamma:

\[ \text{lset rates=gamma shape=0.2 pinvar=0.4;} \]

Note: \textit{estimate}, \textit{previous}, or a specific value can be specified for both \textit{shape} and \textit{pinvar}.
Rate homogeneity in PAUP*

Just tell PAUP* that you want all rates to be equal and that you want all sites to be allowed to vary:

```
lset rates=equal pinvar=0;
```

Note: these are the default settings, but it is useful to know how to go back to rate homogeneity after you have experimented with rate heterogeneity!
Simulating sequences

Exponential probability density

$\exp(2)$

$\exp(5)$

sojourn time

50% of area under blue curve

$\exp(2)$ rate

$\exp(5)$ rate

sojourn time

exp(2)  exp(5)
$p(t) = \lambda e^{-\lambda t}$  

$F(t) = 1 - e^{-\lambda t}$  

Cumulative probability

Random uniform $(0,1)$

$u = 1 - e^{-\lambda t}$

$1 - u = e^{-\lambda t}$

$log(1 - u) = log e^{-\lambda t} = -\lambda t$

$t = -\frac{1}{\lambda} log(1 - u)$
simulate 1 site on 1 edge

\[
\begin{pmatrix}
A & C & G & T \\
-\beta(k + 2) & \beta & \beta k & \beta \\
-\beta(k + 2) & \beta & \beta k & \beta \\
\beta k & \beta & -\beta(k + 2) & \beta \\
\beta & \beta k & \beta & -\beta(k + 2)
\end{pmatrix}
\]

\[
\pi_A = 0.25 < .25 \\
\pi_C = 0.25 < .5 \\
\pi_G = 0.25 < .75 \\
\pi_T = 0.25
\]

\[
\begin{align*}
\text{Subst. rate} &= \beta(k + 2) \\
\text{no. subst.} &= V = \left[\beta(k + 2)\right](t) \\
\text{rate} &= \frac{1}{k + 2} \\
\text{time} &= t
\end{align*}
\]

\[
\begin{align*}
\lambda &= \frac{1}{\beta(k + 2)} \\
\log(1 - u) &= -\frac{1}{\beta(k + 2)}
\end{align*}
\]

\[
\begin{align*}
\log(1 - u) &= \beta(k + 2) \\
\text{measure time in expected no. subst. per site} \\
V &= t
\end{align*}
\]

\[
\begin{align*}
\text{convention:} \\
\text{expected no. subst. per site} \\
\text{simulate 1 site on 1 edge}
\end{align*}
\]
simulate 1 site on tree

\[ t = -\log(1-u) = .28 \]

\[ v = 0.50 \]

\[ v = 0.28 \]

\[ v = 0.21 \]

\[ v = 0.23 \]

\[ v = 0.20 \]

\[ \nu = 0.20 \]

\[ \nu = 0.21 \]

\[ \nu = 0.23 \]

\[ \nu = 0.28 \]

\[ k = 8 \]

\[ \alpha = \frac{1}{10}, \beta = \frac{9}{10} \to T \]

\[ \alpha = \frac{3}{10}, \beta = \frac{7}{10} \to C \]

\[ \alpha = \frac{1}{10}, \beta = \frac{9}{10} \to G \]