Report – asymmetric and displaced inheritance process

May 11, 2022

1 Introduction

Models describing the evolution of phenotypic traits along phylogenies, also called Phylogenetic Comparative Methods (PCM) are crucial to understand the processes that shaped present biodiversity. Most of those models can be described as multivariate stochastic processes derived from the Brownian Motion model (BM) using the phylogeny to describe covariance between species [**?**]. One underlying and often untold hypotheses of those models is the inheritance of the ancestors' phenotype by the descendants at a branching event. This contrasts with a common view in which trait divergence is often involved in pre-zygotic isolation between population leading to speciation. When a trait is involved in the speciation process, it is expected that descendants of the ancestral species will not inherit the same phenotype at the branching event. At the macroevolutionary scale, some continuous traits such as range size, niche width or population size can be asymmetrically inherited between daughter species during the process of speciation. The way these variables are distributed among daughter species is crucial to understand processes undergoing speciation and to identify processes behind the controversial punctuated evolution. This work is an attempt to identify trait evolution with asymmetric inheritance during speciation.

2 model

2.1 Model definition

The goal of this process is to model the evolution of intraspecific phenotypic trait distribution across phylogenies with different scenarios of inheritance at the time of speciation. Specifically, we want to incorporate the situation in which the two incipient species do not fully inherit the ancestral species' phenotypic distribution. The process should allow a continuum between the alternative situations presented in Fig. 1.
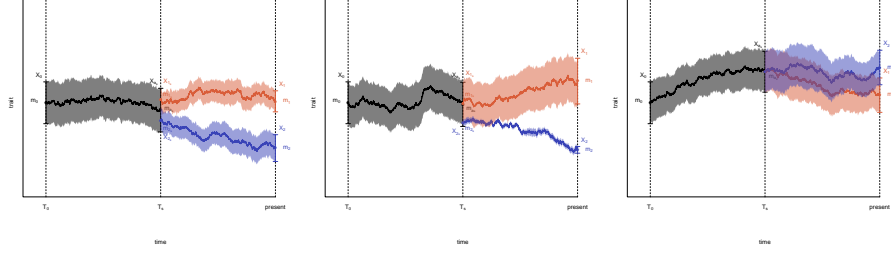
Figure 1: Different scenarios depending on the displacement and asymmetry between the trait distributions of the two descendant species. Left panel: Cladogenetic symmetric inheritance ($\omega = 0.9, \nu = 0$). Center panel: asymmetric inheritance with displacement ($\omega = 0.9, \nu = 0.9$). Right panel: symmetric inheritance (classic BM, $\omega = 0, \nu = 0$)
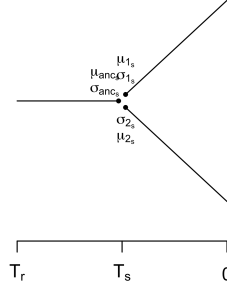

Figure 2: Example of a bifurcation in a phylogenetic tree

We make the hypothesis that the intraspecific trait distribution is normally distributed. Along branches, we model independently the evolution of the intraspecific phenotypic means and variances following standard PCM approaches. At the time of a speciation event (T_s) represented as a bifurcation in a phylogenetic tree (Fig. 2), we first make the hypothesis that the 5th and 95th percentiles of the ancestral distribution are conserved in the union of the descendants' distributions (Fig. 1). Using the inverse error function ($erf^{-1}()$), we can calculate the 5th and 95th of a normal distribution ($X \sim \mathcal{N}(\mu, \sigma)$):

$$\begin{cases} Q_X(0.05) = \mu + \sqrt{(2)}erf^{-1}(2*0.05 - 1)\sigma \\ Q_X(0.95) = \mu + \sqrt{(2)}erf^{-1}(2*0.95 - 1)\sigma \end{cases}$$

Our conservatism hypothesis can then be written as:

$$\begin{cases} Q_a(0.05) = min(Q_{1_s}(0.05), Q_{1_s}(0.05)) \\ Q_a(0.95) = max(Q_{1_s}(0.95), Q_{1_s}(0.95)) \end{cases}$$

2

The inverse error function being an symmetric function with $(erf^{-1}(-0.9) = -erf^{-1}(0.9))$, we can then write:

$$\begin{cases} \mu_a - \Phi^{-1}\sigma_a & = min(\mu_{1_s} - \Phi^{-1}\sigma_{1_s}, \mu_{2_s} - \Phi^{-1}\sigma_{2_s}) \\ \mu_a + \Phi^{-1}\sigma_a & = max(\mu_{1_s} + \Phi^{-1}\sigma_{1_s}, \mu_{2_s} + \Phi^{-1}\sigma_{2_s}) \end{cases} \quad (1)$$

with $\Phi^{-1} = \sqrt{2}erf^{-1}(0.9)$. This equation ensures that the 5th and 95th percentiles of the ancestral distribution are shared respectively with at least one of the descendant distributions. This expression gives the possibility to relax or strengthen the conservatism hypothesis. Second, we postulate that the descendants' distributions might inherit the ancestral variance asymmetrically.

$$\nu = \frac{S_{\nu_1}\sigma_{1_s} + S_{\nu_2}\sigma_{2_s}}{\sigma_a} \quad (2)$$

In this equation, $\nu \in [0,1[$ represents the asymmetry between the descendants' variances and the switch parameters $S_{\nu_1} \in \{-1,1\}$ and $S_{\nu_2} = -S_{\nu_1}$ indicate which descendant inherits the larger part of the ancestral variance. $\nu = 0$ indicates a symmetric inheritance of the ancestral variance while the asymmetry grows when ν comes closer to 1. This equation also adds a second hypothesis about the ancestral distribution's conservation by constraining the difference between descendants' variance to be lower than the ancestral variance. Third, we posit that character displacement might have occurred during speciation that we measure with:

$$\omega = \frac{S_{\omega_1}\mu_{1_s} + S_{\omega_2}\mu_{2_s}}{\Phi^{-1}(2-\nu)\sigma_a} \quad (3)$$

Here $\omega \in [0,1[$ represents the displacement between the descendants' means and the switch parameters $S_{\nu_1} \in \{-1,1\}$ and $S_{\omega_2} = -S_{\omega_1}$ indicate which descendant inherits the highest mean value between the two descendants. When there is no displacement between descendants', ω is equal to 0 while an ω of 1 represents the maximum displacement that also respect the conservatism hypothesis. A third condition of conservation is added here by constraining the displacement to be lower than the 95% interval of the ancestral distribution if the asymmetry is null and half of the 95% interval of the ancestral distribution if the asymmetry is maximal. These equations (1, 2, 3) cover a continuum between alternative inheritance scenarios of the ancestral distribution at the time of speciation (Fig. 3).

2.2 Likelihood of the interspecific distribution

The goal of this section is to estimate the likelihood of the means and variances of observed species given a phylogenetic tree. We start with a simple tree with two species sharing a common ancestor. Using (2) and (3) we can express the variance and mean of the descendants as a function of each other:
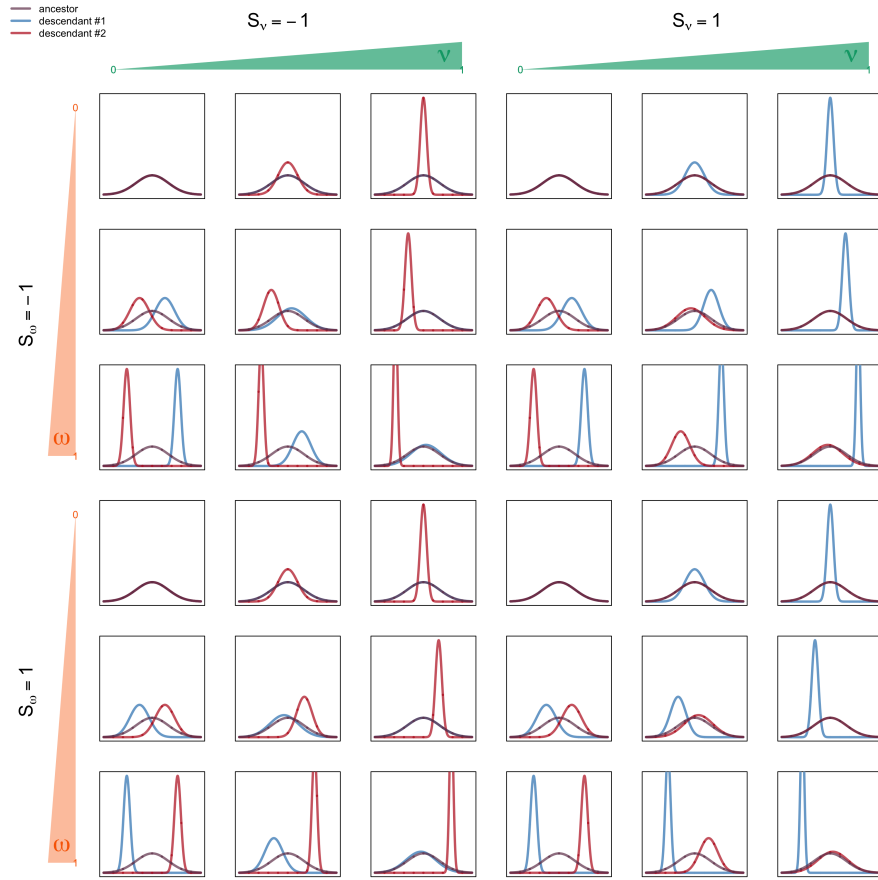
Figure 3: Alternative scenarios covered by the Asymmetric and Displaced Inheritance Process

$$\begin{cases} \sigma_{1_s} &= \sigma_{2_s} - S_{\nu_2}\nu\sigma_a \\ \sigma_{2_s} &= \sigma_{1_s} - S_{\nu_1}\nu\sigma_a \end{cases} \qquad \begin{cases} \mu_{1_s} &= \mu_{2_s} - S_{\omega_2}\omega(2-\nu)\Phi^{-1}\sigma_a \\ \mu_{2_s} &= \mu_{1_s} - S_{\omega_1}\omega(2-\nu)\Phi^{-1}\sigma_a \end{cases}$$

(4)

Combining (4) with (1) we get:

$$\begin{cases} \mu_a - \Phi^{-1}\sigma_a &= min(\mu_{1_s} - \Phi^{-1}\sigma_{1_s}, \mu_{1_s} - \Phi^{-1}\sigma_{1_s} - S_{\omega_1}\omega(2-\nu)\Phi^{-1}\sigma_a + S_{\nu_1}\nu\Phi^{-1}\sigma_a) \\ \mu_a + \Phi^{-1}\sigma_a &= max(\mu_{1_s} + \Phi^{-1}\sigma_{1_s}, \mu_{1_s} + \Phi^{-1}\sigma_{1_s} - S_{\omega_1}\omega(2-\nu)\Phi^{-1}\sigma_a - S_{\nu_1}\nu\Phi^{-1}\sigma_a) \end{cases}$$

which simplifies to:

$$\begin{cases} \mu_a - \Phi^{-1}\sigma_a &= \mu_{1_s} - \Phi^{-1}\sigma_{1_s} + \Phi^{-1}\sigma_a m_1 \\ \mu_a + \Phi^{-1}\sigma_a &= \mu_{1_s} + \Phi^{-1}\sigma_{1_s} - \Phi^{-1}\sigma_a M_1 \end{cases}$$

(5)

With $m1 = min(0, S_{\nu_1}\nu - S_{\omega_1}\omega(2-\nu))$ and $M_1 = min(0, S_{\nu_1}\nu + S_{\omega_1}\omega(2-\nu))$. Solving this set of equations for μ_{1_s} and σ_{1_s} gives us an expression of the first descendant's mean and variance in a function of $\nu, \omega, S_{\nu_1}, S_{\omega_1}$:

$$\begin{cases} \sigma_{1_s} &= \frac{1}{2}\sigma_a(2 + M_1 + m_1) \\ \mu_{1_s} &= \mu_a + \frac{1}{2}\Phi^{-1}\sigma_a(M_1 - m_1) \end{cases}$$

(6)

With the same method for species 2 we get:

$$\begin{cases} \sigma_{2_s} &= \frac{1}{2}\sigma_a(2 + M_2 + m_2) \\ \mu_{2_s} &= \mu_a + \frac{1}{2}\Phi^{-1}\sigma_a(M_2 - m_2) \end{cases}$$

(7)

With $m2 = min(0, S_{\nu_2}\nu - S_{\omega_2}\omega(2-\nu))$ and $M_2 = min(0, S_{\nu_2}\nu + S_{\omega_2}\omega(2-\nu))$.

We consider that species means and variances evolve independently along branches following a Brownian motion. We model the evolution of the logarithm of the standard deviation $(log(\sigma))$ for the variance following the JIVE algorithm; At each node, we apply the asymmetric and displaced inheritance (ADI) process described above. We can calculate the expectations of $log(\sigma_1), log(\sigma_2), \mu_1$ and μ_2 at present according to this model using (6 and 7):

$$\begin{cases} E[log(\sigma_i)] &= log(\sigma_a) - log(2) + log(2 + m_i + M_i) \\ E[\mu_i] &= \mu_a + \frac{1}{2}\Phi^{-1}\sigma_a(M_i - m_i) \end{cases}$$

Even thought it modifies the expectations for our variables, the ADI process does not affect their variances. We thus have standard variances derived from the BM process:

$$\begin{cases} V[log(\sigma_i)] &= \gamma_i\sigma^2_{log(\sigma)} \\ V[\mu_i] &= \gamma_i\sigma^2_\mu \end{cases}$$

γ_i being the length of the terminal branch leading to species i and $\sigma^2_{log(\sigma)}$ and σ^2_μ being respectively the evolutionary rates of $log(\sigma)$ and μ. We can then calculate the probability of $\sigma_1, \sigma_2, \mu_1$ and μ_2 given $\omega, \nu, \sigma_{a_s}, \mu_{a_s}, \sigma^2_\mu$ and $\sigma^2_{log(\sigma)}$ and obtain the likelihood of the tree.
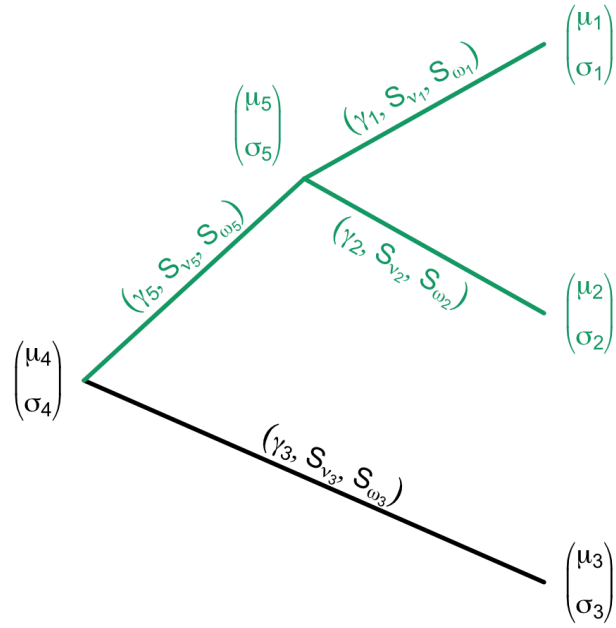
Figure 4: Summary of the notations on the phylogenetic trees: Extant species phenotypic distribution is represented at the tips, ancestral species distribution at the time of speciation is represented at the nodes, branch lengths and inheritance switch parameters are represented along branches. The clade in green represents the initial species complex used to present the ADI process.

To calculate the variances and expectations for a slightly more complicated tree (Fig. 4) we make the hypothesis that ν and ω are constant across every node while we allow every node to have a different value for S_{ν_i} and S_{ω_i} :

$$E[log(\sigma_1)] = E[log(\sigma_5)] - log(2) + log(2 + m_1 + M_1)$$

with

$$E[log(\sigma_5)] = log(\sigma_4) - log(2) + log(2 + m_5 + M_5)$$

leading to

$$E[log(\sigma_1)] = log(\sigma_4) - 2log(2) + log(2 + m_5 + M_5) + log(2 + m_1 + M_1)$$

Similarly we have

$$\begin{cases} E[log(\sigma_2)] = & log(\sigma_4) - 2log(2) + log(2 + m_5 + M_5) + log(2 + m_2 + M_2) \\ E[log(\sigma_3)] = & log(\sigma_4) - log(2) + log(2 + m_3 - M_3) \end{cases}$$

As mentioned earlier, the variances are not affected by the ADIP and are the standard variances of a BM process:

$$\begin{cases} V[log(\sigma_1)] = & (\gamma_5 + \gamma_1)\sigma^2_{log(\sigma} V[log(\sigma_2)] = & (\gamma_5 + \gamma_2)\sigma^2_{log(\sigma} \\ V[log(\sigma_3)] = & \gamma_3\sigma^2_{log(\sigma} \\ Cov[log(\sigma_1), log(\sigma_2)] = & \gamma_5\sigma^2_{log(\sigma} \\ Cov[log(\sigma_1), log(\sigma_3)] = & 0 \\ Cov[log(\sigma_2), log(\sigma_3)] = & 0 \end{cases}$$

Following the same method, we can find expressions $E[\mu_i]$ as a function of μ_4, ω, ν and S_ν, S_ω for every node.

$$\begin{cases} E[\mu_1] = & \mu_4 + \frac{1}{2}\Phi^{-1}\sigma_4(M_5 - m_5) + \frac{1}{2}\Phi^{-1}\sigma_5(M_1 - m_1) \\ E[\mu_2] = & \mu_4 + \frac{1}{2}\Phi^{-1}\sigma_4(M_5 - m_5) + \frac{1}{2}\Phi^{-1}\sigma_5(M_2 - m_2) \\ E[\mu_1] = & \mu_4 + \frac{1}{2}\Phi^{-1}\sigma_4(M_3 - m_3) \end{cases}$$

and

$$\begin{cases} V[\mu_1] = & (\gamma_5 + \gamma_1)\sigma^2_\mu \\ V[\mu_2] = & (\gamma_5 + \gamma_2)\sigma^2_\mu \\ V[\mu_3] = & \gamma_3\sigma^2_\mu \\ Cov[\mu_1, \mu_2] = & \gamma_5\sigma^2_\mu \\ Cov[\mu_1, \mu_3] = & 0 \\ Cov[\mu_2, \mu_3] = & 0 \end{cases}$$

Now if we consider a dichotomous phylogenetic tree with n extant species and $n - 1$ ancestral species all characterized by their trait distribution $X_i \sim \mathcal{N}(\mu_i, \sigma_i)$. Every extant and ancestral species, except the root species, is also characterized by switch parameters $(S_{\nu_i}, S_{\omega_i})$ that controls how it inherited its ancestral distribution at the time speciation and by the length of its ascending branch γ_j (Fig. 4).

By applying (6) and (7) to each node and a BM process to each branch we get :

$$E[log(\sigma_i)] = log(\sigma_{root}) - Jlog(2) + \sum_{j=1}^{J} log(2 + M_j + m_j) \qquad (8)$$

$$E[\mu_i] = \mu_{root} - \frac{1}{2}\Phi^{-1}\sum_{j=1}^{J} \sigma_{a_j}(M_j - m_j) \qquad (9)$$

With J being the number of branches between the root and species i and σ_{a_j} being the standard deviation of the direct ancestor of j at the time of speciation. However, with fixed σ_i, S_{ν_i} and S_{ω_i}, the variance of $\boldsymbol{\mu}$ and $log(\boldsymbol{\sigma})$ remain constant across nodes, allowing the use of a standard phylogenetic variance-covariance matrix. Using $E[\mu_i], E[\sigma_i], V[\mu_i]$ and $V[\sigma_i]$ we can calculate the likelihood functions of $\boldsymbol{\mu}$ and $log(\boldsymbol{\sigma})$ as multivariate normal distributions.

$$\mathcal{L}(\theta|\nu, \omega, \sigma_\mu^2, \sigma_{log(\sigma)}^2, \boldsymbol{\mu'}, \boldsymbol{\sigma'}, \boldsymbol{S_\nu}, \boldsymbol{S_\omega}, \boldsymbol{\gamma})$$
$$\propto P(log(\boldsymbol{\sigma})|\nu, \omega, \sigma_\mu^2, \sigma_{log(\sigma)}^2, log(\boldsymbol{\sigma'}), \boldsymbol{S_\nu}, \boldsymbol{S_\omega}, \boldsymbol{\gamma})$$
$$\times P(\boldsymbol{\mu}|\nu, \omega, \sigma_\mu^2, \sigma_{log(\sigma)}^2, \boldsymbol{\mu'}, log(\boldsymbol{\sigma'}), \boldsymbol{S_\nu}, \boldsymbol{S_\omega}, \boldsymbol{\gamma})$$

With $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$, $\boldsymbol{\mu'}$ and $\boldsymbol{\sigma'}$ being respectively observed means and standard deviations and ancestral means and standard deviations.

2.3   Parameter estimation

The Bayesian estimation of the asymmetry (ν), the displacement (ω) and the evolutionary rates ($\sigma_\mu^2, \sigma_{log(\sigma)}^2$) is dependant on the approximation of ancestral states ($\boldsymbol{\sigma'}, \boldsymbol{\mu'}$) and switch parameters ($\boldsymbol{S_\nu}, \boldsymbol{S_\omega}$). These parameters can be estimated using Gibbs sampling by calculating their conditional distributions on ν, ω, σ_μ^2 and $\sigma_{log(\sigma)}^2$. For our example highlighted in green (Fig. 4), in order to get the conditional distribution of $log(\sigma_5)$ on $log(\sigma_1), log(\sigma_2), \nu$ and ω we can add the expressions from (6) and (7) :

$$log(\sigma_1) + log(\sigma_2) = 2[log(\sigma_5) - log(2)] + log(2 + M_1 + m_1) + log(2 + M_2 + m_2)$$
$$log(\sigma_5) = \frac{1}{2}(log(\sigma_1) + log(\sigma_2)) + C$$

with

$$C = log(2) - \frac{1}{2}[log(2 + M_1 + m_1) + log(2 + M_2 + m_2)]$$

We can show numerically that C does not vary according to S_{ν_1} and S_{ω_1} and therefore that C is constant for a fixed ν and ω across all nodes and can be used as a constant.

$$log(\sigma_4) = \frac{1}{2}(log(\sigma_5) + log(\sigma_3)) + C$$

$$log(\sigma_4) = \frac{1}{2}(\frac{1}{2}(log(\sigma_1) + log(\sigma_2)) + C + log(\sigma_3)) + C$$

$$log(\sigma_4) = \frac{1}{4}(log(\sigma_1) + log(\sigma_2)) + \frac{1}{2}log(\sigma_2) + \frac{1}{2}C + C$$

We can then write $log(\sigma_5)$ and $log(\sigma_4)$ as a linear combination of normal distributions:

$$log(\sigma_5) \sim \mathcal{N}(\frac{1}{2}(E[log(\sigma_1)] + E[log(\sigma_2)]) + C, \frac{1}{4}(V[log(\sigma_1)] + V[log(\sigma_2)]))$$

$$log(\sigma_4) \sim \mathcal{N}(\frac{1}{4}(E[log(\sigma_1)] + E[log(\sigma_2)]) + \frac{1}{2}E[log(\sigma_3)] + \frac{3}{2}C,$$
$$\frac{1}{16}(V[log(\sigma_1)] + V[log(\sigma_2)]) + \frac{1}{4}V[log(\sigma_3)])$$

Using it we can show that for any node k, $log(\sigma_k)$ is a linear combination of its descendants' $log(\sigma)$. Therefore for every node k with I_k extant descendants and J_k descending edges, we have:

$$log(\sigma_k) \sim \mathcal{N}(E[log(\sigma_k)], V[log(\sigma_k)])$$

$$E[log(\sigma_k)] = \sum_{i=1}^{I_k} \frac{1}{2^{n_i}}(E[log(\sigma_i)] + n_i C)$$

$$V[log(\sigma_k)] = \sum_{j=1}^{J_k} \frac{1}{4^{n_j}} \gamma_j \sigma_{log(\sigma)}^2$$

(10)

with n_i and n_j respectively the number of nodes between the node k and i, j. For every node with direct descendants a and b we can also write:

$$P(S_{\nu_a} = 1) \propto P(log(\sigma_a) - log(\sigma_b) > 0)$$
$$X_i = log(\sigma_a) - log(\sigma_b)$$
$$X_i \sim \mathcal{N}(E[log(\sigma_a)] - E[log(\sigma_b)], V[log(\sigma_a)] + V[log(\sigma_b)])$$

(11)

After the sampling of $log(\boldsymbol{\sigma'})$ and $\boldsymbol{S_\nu}$ in their respective conditional distributions, we can calculate the conditional distribution of $\boldsymbol{\mu'}$ and $\boldsymbol{S_\omega}$. In our example, we want to calculate the conditional distribution of μ_5 on $\mu_1, \mu_2, \nu, \omega, S_{\nu_1}$ and S_{ν_2}. In order to get it we want to write an expression of μ_5 as a weighted mean of μ_1 and μ_2.

$$\mu_5 = a_1\mu_1 + a_2\mu_2$$

with $a + b = 1$. Using (6) and (7), we get:

$$\mu_5 = a_1(\mu_5 + \frac{1}{2}\Phi^{-1}\sigma_5(M_1 - m_1)) + a_2(\mu_5 + \frac{1}{2}\Phi^{-1}\sigma_5(M_2 - m_2))$$

$$\mu_5 = (a_1 + a_2)\mu_5 + \frac{1}{2}\Phi^{-1}(a_1(M_1 - m_1) + a_2(M_2 - m_2))$$

$$0 = a_1(M_1 - m_1) + a_2(M_2 - m_2)$$

We define :

$$\begin{cases} l_1 = m_2 = min(0, S_{\omega_2}\omega(2 - \nu) - S_{\nu_1}\nu) \\ l_2 = m_1 = min(0, S_{\omega_2}\omega(2 - \nu) - S_{\nu_2}\nu) \\ L_1 = -M_2 = max(0, S_{\omega_1}\omega(2 - \nu) + S_{\nu_1}\nu) \\ L_2 = -M_1 = max(0, S_{\omega_2}\omega(2 - \nu) + S_{\nu_2}\nu) \end{cases}$$

Then we get:

$$\begin{cases} a_1(M_1 - m_1) + (1 - a_1)(-l_1 - L_1) = 0 \\ (1 - a_2)(-l_2 - L_2) + a_2(M_1 - m_1) = 0 \end{cases}$$

Which simplifies to :

$$\begin{cases} a_1 = \frac{l_1 + L_1}{M_1 - m_1 + l_1 + L_1} \\ a_2 = \frac{l_2 + L_2}{M_2 - m_2 + l_2 + L_2} \end{cases}$$

for $\nu > 0$ and $\omega > 0$. For $\omega = 0$ or $\nu = 0$ the descendants' means or variances variances are equal so we have $a_1 = a_2 = \frac{1}{2}$. We then have:

$$\begin{cases} \mu_5 = \frac{l_1 + L_1}{M_1 - m_1 + l_1 + L_1}\mu_1 + \frac{l_2 + L_2}{M_2 - m_2 + l_2 + L_2}\mu_2 & \text{if} \quad \nu > 0 \quad \text{and} \quad \omega > 0 \\ \mu_5 = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 & \text{if} \quad \nu = 0 \quad \text{or} \quad \omega = 0 \end{cases}$$

We can show that a_1 and a_2 are constant for a fixed ν, ω and S_{ν_1}, so we can calculate the conditional distribution of μ_5 as a linear combination of μ_1 and $\mu2$ which are normally distributed:

$$\mu_5 \sim \mathcal{N}(E[\mu_5], V[\mu_5])$$

$$E[\mu_5] = a_1 E[\mu_1] + a_2 E[\mu_2]$$

$$V[\mu_5] = a_1^2 V[\mu_1] + a_2^2 V[\mu_2]$$

Similarly, for μ_4 we find:

$$\mu_4 = a_5\mu_5 + a_3\mu_3$$

$$\mu_4 = a_5 a_1 \mu_1 + a_5 a_2 \mu_2 + a_3 \mu_3$$

We can show that for any node k, μ_k is a linear combination of its descendants' μ. Therefore for any node k with I_k descendants and J_k descending branches, we have:

10

$$\mu_k \sim \mathcal{N}(E[\mu_k], V[\mu_k])$$

$$E[\mu_k] = \sum_{i=1}^{I_k} E[\mu_i] \prod_{j}^{J_k} a_j \tag{12}$$

$$V[\mu_k] = \sum_{j=1}^{J_k} a_j^2 \gamma_j \sigma_\mu^2$$

For every node with direct descendants a and b we can also write:

$$P(S_{\omega_a} = 1) \propto P(\mu_a - \mu_b > 0)$$

$$Y_a = \mu_a + \mu_b \tag{13}$$

$$Y_a \sim \mathcal{N}(E[\mu_b] - E[\mu_b], V[\mu_a] + V[\mu_b])$$

2.4 model testing