Publication bias and model comparison

Controling for publication bias

The logic for coding the NA is as follows: Given that the true state is a 0, the probability of observing a 0 is $(1 - \psi_0)$, and hence, the probability of **not reporting** such discovery is $(1 - \eta_0)$. Likewise, given a true state equal to 0, the probability of observing a 1 is ψ_0, which has a probability of not reporting equal to $(1 - \eta_1)$. Hence we obtain $\Pr(\text{NA}|\text{True} = 0) = (1 - \eta_0)(1 - \psi_0) + (1 - \eta_1)\psi_0$. The following table list what would be the leaf probabilities for each combination of observed and true state:

True State	0	NA	1
0	$\eta_0(1 - \psi_0)$	$(1 - \eta_0)(1 - \psi_0) + (1 - \eta_1)\psi_0$	$\eta_1\psi_0$
1	$\eta_0\psi_1$	$(1 - \eta_1)(1 - \psi_1) + (1 - \eta_0)\psi_1$	$\eta_1(1 - \psi_1)$

Table 1: How P Thomas suggests including η: labeling probabilities based on the OBSERVED state

Comparing the models with and without η

In computational terms, we can still use the likelihood function that uses η to compute the original model without it. By setting $\eta_0 = \eta_1 = 0.5$, the likelihood of the model will be proportional to the model in which we don't use this parameter by $(1/2)^{P\times|\text{Leafs}|}$, this implies that we can always return to the *baseline* model without the *publication bias parameter*.

$$\text{L}\left(\{\psi, \mu, \eta_0 = \eta_1 = 0.5, \pi\}|X\right) = \text{L}\left(\{\psi, \mu, \pi\}|X\right) \times \left(\frac{1}{2}\right)^{P\times|\text{leafs}|}$$

Such correction needs only to be done if we wish to compare both likelihoods (in for example a likelihood ratio test)[1], as in the Metropolis-Hastings algorithm the constant is cancelled out (we are using symmetric jumps, so we don't need to worry about transition probabilities).

Comparing models with and without ψ

To obtain the likelihood of the model without the ψ parameter, from the computational point of view, it suffices to set $\psi_0 = \psi_1 = 0$ in the likelihood function to have a model in which mislabeling plays no role. Formally:

$$\text{L}\left(\{\psi_0 = \psi_1 = 0, \mu, \pi\}|X\right) = \text{L}\left(\{\mu, \pi\}|X\right)$$

[1] The latest version of `aphylo` includes a test that checks this statement, this is, whether the likelihood of the model with the eta parameters equal to 0.5 is proportional to the likelihood with no eta parameters at all.