In the past we've run into this problem when we exponentiate large negative LLik values, such as when we need to do a summation of the probability of the set of possible outcomes.

For example, imagine $p_i = \exp[\ell_i]/(\sum_j \exp[\ell_j])$. If the ℓ values are all < -312, exponentiating these terms will lead to 0s and the sum of these 0s will also be 0.

This is a numerical issue, not a fundamental problem with the model. The best solution for this is to define

$$\ell_{\max} = \max(\vec{\ell}) \tag{0.1}$$

and then rescale ℓ terms by subtracting off $\ell_{\max}$. Let ℓ'_i represent the rescaled terms, such that $\ell'_i = \ell_i - \ell_{\max}$. They key result is that one of these terms ℓ'_i term will always evaluate to 1 when exponentiated and the rest will always evaluate to < 1 when exponentiated.

Why does this work? Well, subtracting a constant term from all the exponents is equvalent to dividing all of them by $\exp[\ell_{\max}]$ and if we apply this division to both the numerator and denominator terms of p_i, its value will be unchanged (but evaluatable). That is,

$p_i = \exp[\ell_i]/(\sum_j \exp[\ell_j]) = \exp[\ell'_i]/(\sum_j \exp[\ell'_j])$

I hope this helps.

Dr. Gilchrist