SOME HARMONIC NUMBER IDENTITIES FOR PHYLOGENETIC TREE ANALYSIS

Krzysztof Bartoszek

Linköping University

e-mail: krzysztof.bartoszek@liu.se, krzbar@protonmail.ch

Values associated with phylogenetic trees like the total tree area [4] or the cophenetic index [5] can be represented through the height of the tree, and the time to coalescent of a random pair of tips. In this way the given index, for a random tree, can be studied by considering a pair of (dependent) one–dimensional random variables. Control over their moments will immediately provide information on the behaviour of these indices [1, 3, 6]. In order to obtain these moments, for the pure birth tree, one has to consider rather involved harmonic and quadratic harmonic sums. In the finite term case, these sums often turn out to have closed form formulæ in terms of harmonic numbers. However, surprisingly, symbolic algebra systems do not seem to be able (at least out of the box) to find these final forms. In our talk we will show how these sums arrive in the analysis of tree heights, in what situations computer algebra systems fail, and how one can approach these sums [2].

References

[1] K. Bartoszek, Quantifying the effects of anagenetic and cladogenetic evolution. Mathematical Biosciences 254 (2014) 42–57.

[2] K. Bartoszek, Closed and asymptotic formulæ for harmonic and quadratic harmonic sums. ArXiv e-prints (2023). arXiv: 2312.15366.

[3] K. Bartoszek, Exact and approximate limit behaviour of the Yule tree's cophenetic index. Mathematical Biosciences 303 (2018), 26–45.

[4] A. Mir, F. Rosselló, The mean value of the squared path–difference distance for rooted phylogenetic trees. Journal of Mathematical Analysis and Applications 371 (2010), 168–176.

[5] A. Mir, F. Rosselló and L. Rotger, A new balance index for phylogenetic trees. Mathematical Biosciences 241 (2013), 125–136.

[6] S. Sagitov, K. Bartoszek, Interspecies correlation for neutrally evolving traits. Journal of Theoretical Biology. 309 (2012) 11–19.