Ann Inst Stat Math (2018) 70:99–129
https://doi.org/10.1007/s10463-016-0582-9

Limiting behaviour of Fréchet means in the space of phylogenetic trees

D. Barden[1] · H. Le[2] · M. Owen[3]

Received: 3 March 2016 / Revised: 31 August 2016 / Published online: 19 October 2016

Abstract As demonstrated in our previous work on $\boldsymbol{T}_4$, the space of phylogenetic trees with four leaves, the topological structure of the space plays an important role in the non-classical limiting behaviour of the sample Fréchet means in $\boldsymbol{T}_4$. Nevertheless, the techniques used in that paper cannot be adapted to analyse Fréchet means in the space $\boldsymbol{T}_m$ of phylogenetic trees with $m(\geqslant 5)$ leaves. To investigate the latter, this paper first studies the log map of $\boldsymbol{T}_m$. Then, in terms of a modified version of this map, we characterise Fréchet means in $\boldsymbol{T}_m$ that lie in top-dimensional or co-dimension one strata. We derive the limiting distributions for the corresponding sample Fréchet means, generalising our previous results. In particular, the results show that, although they are related to the Gaussian distribution, the forms taken by the limiting distributions depend on the co-dimensions of the strata in which the Fréchet means lie.

Keywords Central limit theorem · Fréchet mean · Log map · Phylogenetic trees · Stratified manifold

✉ H. Le
huiling.le@nottingham.ac.uk

D. Barden
d.barden@dpmms.cam.ac.uk

M. Owen
megan.owen@lehman.cuny.edu

1 Girton College, University of Cambridge, Cambridge CB3 0JG, UK

2 School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK

3 Department of Mathematics and Computer Science, Lehman College, City University of New York, Bronx 10468, USA