



Statistics and Actuarial Science Seminar

Title: Manifold valued data analysis of samples of networks, with applications in corpus linguistics

Speaker: Katie Severn (University of Nottingham)

Date: Mon 22nd March 2021 at 12:00PM

Location: Online

Abstract: Networks can be used to represent many systems such as text documents and brain activity, and it is of interest to develop statistical techniques to compare networks. We develop a general framework for extrinsic statistical analysis of samples of networks, motivated by networks representing text documents in corpus linguistics. We identify networks with their graph Laplacian matrices, for which we define metrics, embeddings, tangent spaces, and a projection from Euclidean space to the space of graph Laplacians. This framework provides a way of computing means, performing principal component analysis and regression, and performing hypothesis tests, such as for testing for equality of means between two samples of networks. We apply the methodology to the set of novels by Jane Austen and Charles Dickens.

Join Zoom Meeting: <https://ucd-ie.zoom.us/j/68316324831>