The Annals of Applied Statistics
2009, Vol. 3, No. 1, 458–488
DOI: 10.1214/08-AOAS206

MULTILEVEL FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS[1]

BY CHONG-ZHI DI, CIPRIAN M. CRAINICEANU, BRIAN S. CAFFO AND NARESH M. PUNJABI

Johns Hopkins University

The Sleep Heart Health Study (SHHS) is a comprehensive landmark study of sleep and its impacts on health outcomes. A primary metric of the SHHS is the in-home polysomnogram, which includes two electroencephalographic (EEG) channels for each subject, at two visits. The volume and importance of this data presents enormous challenges for analysis. To address these challenges, we introduce multilevel functional principal component analysis (MFPCA), a novel statistical methodology designed to extract core intra- and inter-subject geometric components of multilevel functional data. Though motivated by the SHHS, the proposed methodology is generally applicable, with potential relevance to many modern scientific studies of hierarchical or longitudinal functional outcomes. Notably, using MFPCA, we identify and quantify associations between EEG activity during sleep and adverse cardiovascular outcomes.

1. Introduction.

1.1. *Data description.* The Sleep Heart Health Study (SHHS) is a large-scale comprehensive multi-site study of sleep and its correlation with health outcomes. In the following section we provide a detailed description of the study, and summarize some organizational and demographic characteristics. The principal aim of the study is to learn about the association between sleep and a variety of health-related conditions. The study is specifically designed to examine the potential associations between sleep-disordered breathing (SDB) and outcomes such as hypertension and cardiovascular disease (CVD) [Quan et al. (1997)]. In our example analysis we focus on hypertension, a proposed consequence of disturbed sleep [Shahar et al. (2001)].

We now present a summary of the SHHS characteristics and our scientific hypotheses. A more detailed description of the SHHS can be found in Quan et al. (1997) and Crainiceanu et al. (2009). The SHHS is a multi-center study that utilized the resources of existing, well characterized, epidemiologic cohorts, and conducted further data collection, including measurements of sleep and breathing.

Received May 2008; revised September 2008.

[1]Supported by Grants R01NS060910 from the National Institute of Neurological Disorders and Stroke, HL083640, HL07578, and AG025553 from the National Heart, Lung, and Blood Institute and K25EB003491 from the National Institute of Biomedical Imaging and BioEngineering.

Key words and phrases. Functional principal component analysis (FPCA), multilevel models.