BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Sewell, Daniel K

eRA COMMONS USER NAME (credential, e.g., agency login): DKSEWELL

POSITION TITLE: Assistant Professor

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Harding University	B.A.	05/2006	Education
University of Arkansas	M.S.	05/2010	Statistics
University of Illinois Urbana-Champaign	Ph.D.	05/2015	Statistics

A. Personal Statement

I am an Assistant Professor of Biostatistics at the University of Iowa and a faculty member of the Iowa Informatics Initiative, with research interests in network analysis, clustering, Bayesian methodology, statistical computation, and data visualization. The primary focus of my research is in the development of statistical models and inference for network data, and in applying network methodology to answer important scientific questions. I have published papers on this topic in top statistical journals. I have collaborated with researchers from a myriad of scientific disciplines through my role as a statistical consultant with the Department of Statistics at the University of Illinois at Urbana-Champaign and through working with other faculty members at the University of Illinois, the University of Iowa, and at other R1 institutions. I have also developed statistical software and web-based apps implementing novel methodology.

a) Sewell, D. K. and Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association* **110**(512), 1646-1657.
b) Sewell, D. K. and Chen, Y. (2015). Analysis of the formation of the structure of social networks using latent space models for ranked dynamic networks. *Journal of the Royal Statistical Society, Series C* **64**, 611-633.
c) Sewell, D. K., and Chen, Y. (2017). Latent space approaches to community detection in dynamic networks. *Bayesian Analysis* **12**(2), 351-377.
d) Sewell, D. K. (2017). Network autocorrelation models with egocentric data. *Social Networks,* **49**, 113-123.

B. Positions and Honors

Positions and Employment

2008-2009	Teaching Assistant (Instructor), Department of Mathematical Sciences, University of Arkansas
2010	Research Assistant, Department of Mathematical Sciences, University of Arkansas
2010-2011	Teaching Assistant (Lon-Capa Programmer), Department of Statistics, University of Illinois Urbana-Champaign
2011-2012	Research Assistant to Professor Ping Ma, University of Illinois Urbana-Champaign

2012-2015	Research Assistant to Professor Yuguo Chen, University of Illinois Urbana-Champaign
2014	Teaching Assistant (Instructor), Department of Statistics, University of Illinois Urbana-Champaign
2015-present	Assistant Professor, Department of Biostatistics, University of Iowa

Honors

2007	Outstanding Calculus Award, Harding University
2014	Graduate College Travel Award, University of Illinois
2014	Finalist for the Norton Prize for Outstanding Doctoral Thesis in Statistics, University of Illinois
2014	Patrick J. Fett award for the best paper on the scientific study of Congress and the presidency
2015	New Faculty Research Award, College of Public Health, University of Iowa
2016	University of Illinois Nominee for GS/ProQuest Distinguished Dissertation Award (limit 1 per university)

C. Contributions to Science.

My research has focused primarily on network analysis, clustering, computational statistics, and applied statistics. I also have extensive consulting and collaborative experience, working with scientists from a wide range of fields. A full listing of my published work can be found at https://www.ncbi.nlm.nih.gov/myncbi/browse/collection/51171056/?sort=date&direction=descending.

Network Analysis: Network analysis is a ubiquitous area of study used by scientists in many distinct fields, and can most simply be thought of as the study of how objects interact or are connected. The statistical analysis of network data is an inherently difficult endeavor, however, and sophisticated tools are necessary to obtain correct inference. My research has focused on developing models and methods for analyzing a variety of types of network data. The first paper develops a framework for modeling temporally measured network data, and provides tools for prediction, handling missing data, and detecting influence throughout the network. The next two papers develop tools for modeling networks constructed from various types of complex interactions or connections; the second paper also develops methods for quantifying and estimating the stability of the network and determining the timing and structure of subgroup formation in temporal networks. The fourth paper provides an analytical framework for analyzing partially observed networks (egocentric network data), appropriately modeling the heteroscedastic and correlated errors involved in network data as well as estimating the effect of social influence on the response. I have also submitted to and received an invitation for a minor revision for *Network Science* a paper describing a parsimonious method of understanding very complex relationships between actors in longitudinal data. Submitted to the *Social Networks* is work I have done on handling heterogeneity in social influence models.

a) Sewell, D. K. and Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association* **110**(512), 1646-1657.
b) Sewell, D. K. and Chen, Y. (2015). Analysis of the formation of the structure of social networks using latent space models for ranked dynamic networks. *Journal of the Royal Statistical Society, Series C* **64**, 611-633.
c) Sewell, D. K. and Chen, Y. (2016). Latent space models for dynamic networks with weighted edges. *Social Networks* **44**, 105-116.
d) Sewell, D. K. (2017). Network autocorrelation models with egocentric data. *Social Networks,* **49**, 113-123.

Clustering analysis: Numerous applications exist in which it is of interest to partition the data into homogeneous groups, or clusters. This is particularly difficult in longitudinal data, where researchers must think carefully about how to define such clusters within their specific contexts. The first paper derives a statistical model which takes into account temporal dependencies in the data, allows covariate information to help explain the grouping/clustering, and develops computationally efficient algorithms to make model estimation feasible. Clustering within networks, called community detection, is a particularly challenging task. This is even more so when the network data is collected over time. The second paper develops a statistical framework that allows one to perform community detection on temporally collected network data, and again computationally efficient algorithms are derived to make the model estimation feasible. The methods of this has also been published in the R package 'dnc'.

a) Sewell, D. K., Chen, Y., Bernhard, W. and Sulkin, T. (2016). Model-based longitudinal clustering. *Statistica Sinica* **26**(1), 205-233.
b) Sewell, D. K., and Chen, Y. (2016). Latent space approaches to community detection in dynamic networks. Accepted at *Bayesian Analysis*.
c) Sewell, D. K. (2016). dnc: Dynamic Network Clustering (R package). https://cran.r-project.org/web/packages/dnc/index.html

Longitudinal Data Analysis: Studies often require a temporal component in order to track changes over time or to better understand an ongoing process. Failure to correctly account for the additional dependencies in the data induced by this temporal component can lead to biased results, wrong standard errors, and hence wrong inference and conclusions gleaned from the data. Much of my methodological research has focused on temporal, or longitudinal, data analysis. The first paper takes a commonly used network analytic approach and extends it in a non-trivial way to networks observed over time, allowing for the underlying network structure to evolve over time and providing prediction tools for future unobserved time points. The second paper also provides novel methodology for analyzing longitudinal, or dynamic, network data. In this work I provide explicit mechanisms for the estimation of changes over time in the stability and emerging clustering structure of the network. The third paper develops a model-based clustering algorithm for multivariate longitudinal data. The fourth paper considers the temporal variation in surgical site infections (SSIs); in particular, we estimate the seasonality of SSIs and the proportion of that seasonality that can be explained by climate variables.

a) Sewell, D. K. and Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association* **110**(512), 1646-1657.
b) Sewell, D. K. and Chen, Y. (2015). Analysis of the formation of the structure of social networks using latent space models for ranked dynamic networks. *Journal of the Royal Statistical Society, Series C* **64**, 611-633.
c) Sewell, D. K., Chen, Y., Bernhard, W. and Sulkin, T. (2016). Model-based longitudinal clustering. *Statistica Sinica* **26**(1), 205-233.
d) Anthony, C.A., Peterson, R.A., Polgreen, L.A., Sewell, D.K., and Polgreen, P.M. (2017). The seasonal variability in surgical site infections and association with warmer weather: a population-based investigation. *Infection Control & Hospital Epidemiology*. 1-8. doi:10.1017/ice.2017.84.

Computational Statistics: Statistical tools can be implemented to accomplish otherwise intractable computational problems, such as numerical integration or optimization. Network analysis is notorious for having high computational cost. In the first paper, I derive a novel estimation algorithm that reduces the computational cost from exponential to linear in the number of time points, thereby making an otherwise infeasible estimation procedure have extremely short run times. In the second paper, I implement a variational Bayes approach to perform community detection, i.e., clustering, on dynamic networks, reducing the computational burden by around 95%. In the third and fourth papers, I extend a method of reducing computational cost from quadratic to linear in the number of actors in the network to longitudinal network data and to a wide range of distributions, namely any exponential family of distributions. Additionally, I have two as of yet unpublished papers contributing to statistical computing. The first derives a method with which to perform on-line learning of static parameters in a general state space framework through a novel sequential Monte Carlo algorithm. The second provides a new way of deriving, and thus understanding, a well-known estimator for the marginal likelihood, which is often an intractable integral key to many model selection approaches.

a) Sewell, D. K., Chen, Y., Bernhard, W. and Sulkin, T. (2016). Model-based longitudinal clustering. *Statistica Sinica* **26**(1), 205-233.
b) Sewell, D. K., and Chen, Y. (2016). Latent space approaches to community detection in dynamic networks. Accepted at *Bayesian Analysis*.
c) Sewell, D. K. and Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association* **110**(512), 1646-1657.
d) Sewell, D. K. and Chen, Y. (2016). Latent space models for dynamic networks with weighted edges. *Social Networks* **44**, 105-116.

Applied Statistics: Statistical analyses are crucial to most empirical studies, and without these analytical techniques it would not be possible to make meaningful and reliable conclusions from data and to understand how confident we ought to be in these conclusions. I have worked with a myriad of researchers across multiple disciplines. This began as a statistical consultant, working for the Department of Statistics at the University of Illinois at Urbana-Champaign, in which I worked on dozens of research projects. It has continued as I have connected with various academic researchers, leading to interesting methodological and substantive results. Below is a sample of some published work from these collaborative efforts.

a) Sewell, D. K., Kim, H., Ha, T. and Ma, P. (2015). A parameter estimation method for single molecule fluorescence lifetime data. *BMC Research Note* **8**(1), 230.
b) Bernhard, W., Sulkin, T., and Sewell, D. K. (2016). Explaining legislative style in the House of Representatives. Invited revision to *Legislative Studies Quarterly.*
c) Metcalf, K., Baquero, B., Laroche, H., Garcia, M., Francis, S., Janz, K., and Sewell, D.K. (2017+) Calibration of the global physical activity questionnaire to accelerometry measured physical activity and sedentary behavior in a random sample of rural Iowans. Submitted to *International Journal of Behavioral Nutrition and Physical Activity*.
d) Anthony, C.A., Peterson, R.A., Polgreen, L.A., Sewell, D.K., and Polgreen, P.M. (2017). The seasonal variability in surgical site infections and association with warmer weather: a population-based investigation. *Infection Control & Hospital Epidemiology*. 1-8. doi:10.1017/ice.2017.84.

D. Research support

1. RFA-CK-17-001 — 08/01/2017-07/31/2020
   PI:Alberto Segre
   Co-I: Daniel K. Sewell

2. University of Iowa College of Public Health Opportunity Award — 05/01/2017-12/31/2017

3. University of Iowa College of Public Health New Faculty Research Award — 01/01/2016-12/31/2016

4. eHealth Project 82154300 — 01/21/2016-01/21/2017
   PI: Polgreen, Philip

5. U48 DP001902 — 01/26/2016-01/26/2017
   PI: Parker, Edith
   co-PI: Baquero, Barbara