Bayesian Inference in Ecological Research in the Water Resource Evaluation (WRE) Section
DR-14-03; July 2014

Abel Porras, P.E., Aaron Richter, E.I.T
Water Resource Evaluation Section
City of Austin Watershed Protection Department

Introduction

In of support of the City of Austin Watershed Protection Department mission to protect water quality, the Water Resource Evaluation Section (WRE) frequently conducts scientific studies evaluating anthropogenic impacts on various hydrological, ecological and biological regimes. These studies often culminate in white paper reports that could be published in scientific journals. To protect the quality and integrity of the research, a Quality Assurance Project Plan (QAPP) is generated which documents the planning process of the study, provides the objectives of the study in a concise and complete format, and describes the generation, validation and assessment of the data. Projects passing through this process allow that the data collected from the studies is stored securely and accurately in the WRE Field Sampling Database (FSDB).

One difficulty with implementing the QAPP process is for the authors of a QAPP to agree upon a sample design that is believed to generate data in such a way that it may be properly assessed, is flexible to change over time and will not be overly taxing on staff. The latter conditions often can put stress on the first condition when the data is analyzed under the traditional *frequentist* approach, which has been the most commonly practiced form of statistics for many years. The *frequentist* approach results in a qualitative conclusion based on the probability of the data or more extreme data occurring given a hypothesis (McCarthy 2007). This is commonly referred to as null hypothesis testing and in the past twenty years has faced increasing dissent among scholars (Berger and Berry 1988; Johnson 1999; Bakan 1966; Anderson et al. 2000) for its misuse in ecological studies.

Ecological studies involving large datasets are particularly susceptible to these foundational issues. Additionally, WRE QAPPs have, in the past, been conceived of as “living documents” that are open to documented changes, and gives the impression that every incipient change is an allowable one. Ending the experiment, switching to a different mix of treatments, ceasing to admit certain types of subjects or dropping an experimental site, all affect the conclusion of the study when operating under a *frequentist* approach and exacerbates the misuse of the *frequentist* theory.

Post-hoc analyses are sometimes performed on the collected data set, which is again not in accordance with a *frequentist* philosophy. Typically, with projects that have accumulated large amounts of data, every measurement is thrown into every conceivable (and linear) model to determine whether any of those parameters can be used as a covariate. This kind of data mining (also called “fishing”, which is a form of data dredging) is inappropriate in most ecological settings and specifically when the initial assumption of data collection was a more specific oriented hypothesis test, as dictated by a QAPP. (Selvin and Stuart 1966).

Other forms of statistics that have recently become more widely used are Bayesian statistics and Machine Learning. These two forms of statistics are becoming increasingly popular due to the increase in computer power as they are computationally intensive processes. The Bayesian approach calls for conclusions to be made based on the likelihood of the hypothesis fitting the data, rather than the likelihood of the data fitting a hypothesis (which is sometimes known beforehand to be wrong). Conclusions obtained from the Bayesian approach are not affected by the fact that the data are continuously being monitored or changed (Berger and Berry 1988), can be used as an alternative to the common frequentist's statistical tests and can be used in post-hoc analysis (Waller and Duncan 1969; Dempster et al. 1977).

Additionally, there is a growing body of scientific literature which argues for Bayesian inferences, especially for use in ecology where replication and randomization of the samples are difficult to manage. For instance, Bayesian inference does not require large sample sizes to make appropriate inferences, and it has the advantage of providing an intuitive and natural interpretation of probability in its results, which is consistent with how laypeople interpret probabilities in *frequentist* statistics. Abiding by just a *frequentist* approach in the face of an expansive scholarship would appear to be imprudent and irresponsible to the mission of the department.

Recommendation

As a result of the advantages of Bayesian inference, the authors propose to utilize the Bayesian approach for projects where the scope of the data collection is extensive. As a first approximation (future input may lead to a stronger criteria that is also conducive with other teams within the department), extensive data collection can be interpreted as data sets with either greater than one year of sampling, or greater than 10 sampling sites, or greater than 3 sampling parameters (exclusive of the typical four conventional physiochemical field parameters of dissolved oxygen, water temperature, pH, and conductivity). Additionally, these extensive projects would consist of at least two phases. Phase One would consist of a smaller scale to test the validity of the project prior to the intensive time, effort and resources, typically allocated to one of the larger projects. To continue the tradition of maintaining a living QAPP, the data objectives of Phase Two could be adapted based on the conclusions of Phase One. This could entail using the posterior distribution (i.e., the answer) from Phase One as a prior for Phase Two. This would allow flexibility in the experimental design, an efficient sample size and an appropriate inference, all in accordance with Bayesian Theory. A written report would be done on the analysis of Phase One and then if the method has been shown to be valid, then the larger portion of the project could be done.

Not every project would require this level design. Methodologies which are easily agreed upon, projects which are smaller in scale such that an initial phase would be the entire project or projects in which the speed of the data collection is required by some timeline of an outside entity, would fall outside the phased in process. However, the process can remain flexible enough to encapsulate these types of projects. And, as always, every project would necessitate a QAPP and all data collected for the QAPP would be stored in the FSDB.

Case Study

City staff and volunteers began collecting data on invasive terrestrial plant species and their environment in 2013 under the Austin Invasive Species Management Program. The result of this data collection was information on invasive and native plant species from over 2,000 plots spread over 20 parks with each plot consisting of data from a canopy layer, an understory layer, and a ground cover layer. Even though null hypotheses were specified in the program's QAPP, "fishing" was prevalent during the data analysis phase. Furthermore, since in ecology "everything is correlated with everything else", the null hypotheses were rejected, but diagnostic plots had revealed that these conclusions were inappropriate since assumptions, such as homoscedasticity, had not been met.

Under the approach recommended in this paper, only three or four parks would be sampled under Phase One. This would ameliorate the variability in the data from the different parks. Since Bayesian inference does not require large sample sizes, it can be used to look for general trends in invasive plant environments both within and between parks.

Data analysis using Bayesian inference can be typically implemented in the following manner.

1. Estimate counts of invasive and/or native plants between each layer for each park.
2. Account for overdispersion, zero inflation, and offsets (i.e., accounting for variation in plot size).
3. Within each park, fit various covariates and estimate the parameter coefficients for these covariates.
4. Finally, perform a mixed-effects model to look at covariate estimates as random variables between parks.

The conclusions from this analysis would result in a probability of the different hypotheses being true given the data. And, once these probabilities are assigned to each of the hypotheses in these steps, those hypotheses with the highest probability of occurring given the data can be used for a second round of sampling and testing at other parks in subsequent phases.

Conclusion

In performing Bayesian inferences on certain (if not all) ecological projects, WRE can benefit in three ways. First, ecologically relevant questions such as "What is the probability that a certain hypothesis is true?" and "What is the probability that a certain parameter will take this range of values?" can be addressed in a satisfactory manner. This is in contrast to a *frequentist* who asks "What is the probability that the data fits a prescribed hypothesis?" Second, large sample sizes are not required for Bayesian statistics, which can provide substantial cost benefits. Third, the experience and knowledge of the ecologist/biologist can be initially incorporated into the data analysis in the form of a prior, which is becoming a more scientifically appropriate way of investigating ecological phenomena. The results, experience and knowledge gained from that study can then be carried forward as priors into future studies. This can ensure that the results, experiences, and knowledge are not lost or forgotten and that science in WRE is conducted in a cumulative and effective manner.

References

Anderson D, Burnham K, Thompson W. 2000. Null hypothesis testing: problems, prevalence, and an alternative. The Journal of Wildlife Management. 64:912-923. Available from: http://www.jstor.org/stable/3803199

Bakan, D. 1966. The test of significance in psychological research. Physchological Bulletin 66:423-437.

Berger J., Berry D. 1988. Statistical analysis and the illusion of objectivity. American Scientist 76:159-165.

Dempster A., Schatzoff M., Wermuth N. 1977. A simulation study of alternatives to ordinary least squares. Journal of the American Statistical Association 72:77-91.

Johnson, D.H. 1999. The insignificance of statistical significance testing. The Journal of Wildlife Management 63:763-773.

McCarthy, M. 2007. Bayesian Methods for Ecology. Cambridge University Press. New York.

Selvin H., Stuart A. 1966. Data-Dregding Procedures in Survey Analysis. The American Statistician 20:20-23.

Waller R., Duncan D. 1969. A Bayes rule for the symmetric multiple comparison problem. Journal of the American Statistical Association 64:1484-1503.