William Cipolli
7030 Payne St.
Hamilton, NY 13346
+1 (203) 848 5643
will@cipolli.com
www.cipolli.com

Scholarly Activity

My research interests can be roughly categorized into four areas: Bayesian nonparametrics, statistical collaboration, multidisciplinary collaboration, and statistical education. As of March 2022,I have been an author on 30 papers; 26 are published, 1 is accepted, and 4 are submitted or in revision. I have also collaborated on three successful grants; and I am pursuing new funding opportunities.

This section discusses my active research program and reflects on my future research plans to maintain an active research program; build new collaborative relationships with statisticians and scientists from other areas; and be involved in research activities at the national and international levels.

Bayesian Nonparametrics

Traditional parametric procedures in statistics assume that the data are distributed in a particular way. While some of these procedures work when the assumptions are not met, many do not, and faulty assumptions can lead to incorrect conclusions. As a result, a first step for statisticians is often to check whether a plot of the data closely tracks the assumed distribution. However, what can be done when the data do not match the assumed distribution or follow a known distribution?

A subfield of statistics that deals with this issue is called Bayesian nonparametrics. In this area, I focus on computationally efficient solutions that incorporate Polya trees, which are used to flexibly model the distribution of data. The Polya tree distribution can be conceptualized as the normal distribution with the necessary changes "painted on" to make various distribution shapes possible (Figure 1). In the work I detail below, I have successfully advanced the methodological literature by facilitating data analysis when critical assumptions have not been met.

In Cipolli et al. (2016) I propose a novel approach for researchers who must conduct many statistical tests at once. For example, say a pharmaceutical company considers 10,000 chemical compounds for developing a new drug. Even if none of these compounds has an actual effect, we would expect roughly 500 to pass Phase I testing, 25 to pass Phase II testing, and one to pass Phase III testing due to false positives in statistical experiments. Without corrections to our testing procedures, we are exposed to

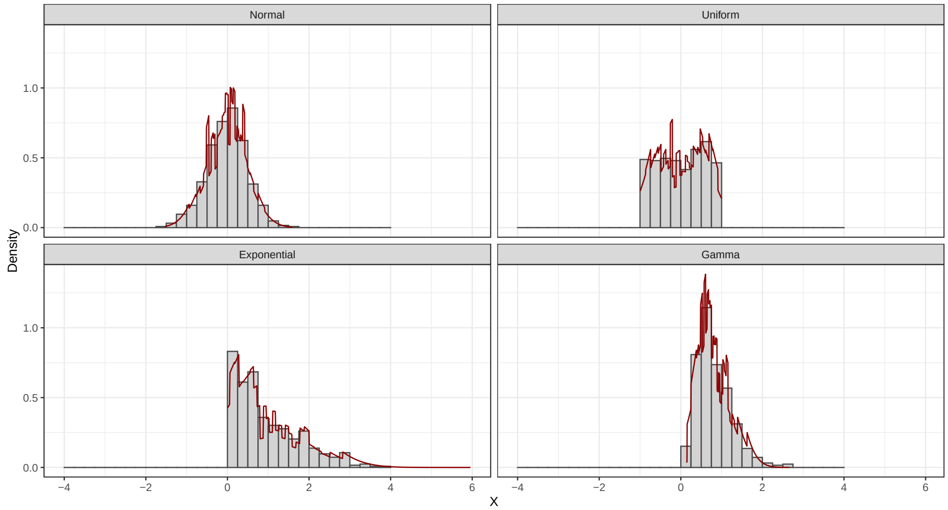

Figure 1: Histograms of samples from four named continuous probability distributions with Polya tree estimates superimposed.

this risk. Scott and Berger (2006) initially proposed a Bayesian approach that assumes the observations are either unchanged (no effect) or shifted (the effects are well-modeled by a normal distribution). The Polya tree framework I developed extends this approach to the Bayesian nonparametric setting and to the heteroskedastic case, where the variance of effects can fluctuate across tests (e.g., compounds). To accomplish this, I developed a discrete approximation to a Polya tree density estimate that can classify unchanged and shifted observations and flexibly estimate the distribution of the shifts. Critically, this computationally fast approach works well when the effects follow the normal distribution *and* when they do not (Figure 2).

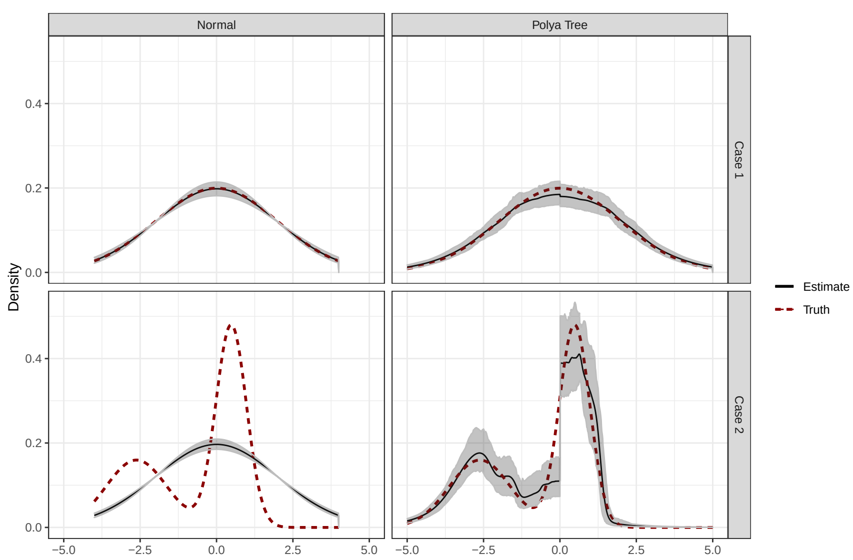

Figure 2: Two examples of multiple testing – case 1 where the shifts are truly normal and case 2 where the shifts follow a bimodal distribution. The average density estimate over 100 simulations is plotted with a 90% confidence band for the Polya tree approach and an approach that assumes the data are normal.

One drawback to this approach is that the Polya tree density estimates are very spiky (Figure 1). To evaluate the long-run ability of these estimation procedures, researchers will simulate data hundreds of times and average the resulting estimates as in Figure 2. While this summary of the estimates closely tracks the actual distribution, the individual estimates experience the same spikiness exhibited in Figure 1. To address this issue, I developed a smoothed Polya Tree distribution that paints the detail with less jagged brush strokes.

In Cipolli and Hanson (2017) I propose computationally efficient methodologies for distribution estimation via two novel nonparametric approaches based on Polya trees. The first formalizes the approach used for multiple testing for density estimation of continuous data, and the second offers a method for smooth Polya tree estimates. Notably, the smoothing technique addresses a commonly raised issue of spikiness in distribution estimates with the standard Polya tree approach.

As an applied example, consider that the distribution of MFAP4 levels is unknown and appears to differ by fibrosis stage (note that these data also appear in the teaching section of this statement in Figure 1). The Polya tree-based distribution results roughly track the histograms despite substantial deviations from normality and provide insight into how MFAP4 levels vary depending on disease status without assuming the distribution *a priori* (Figure 3). Further, the smoothed Polya tree estimates are significantly less spiky.

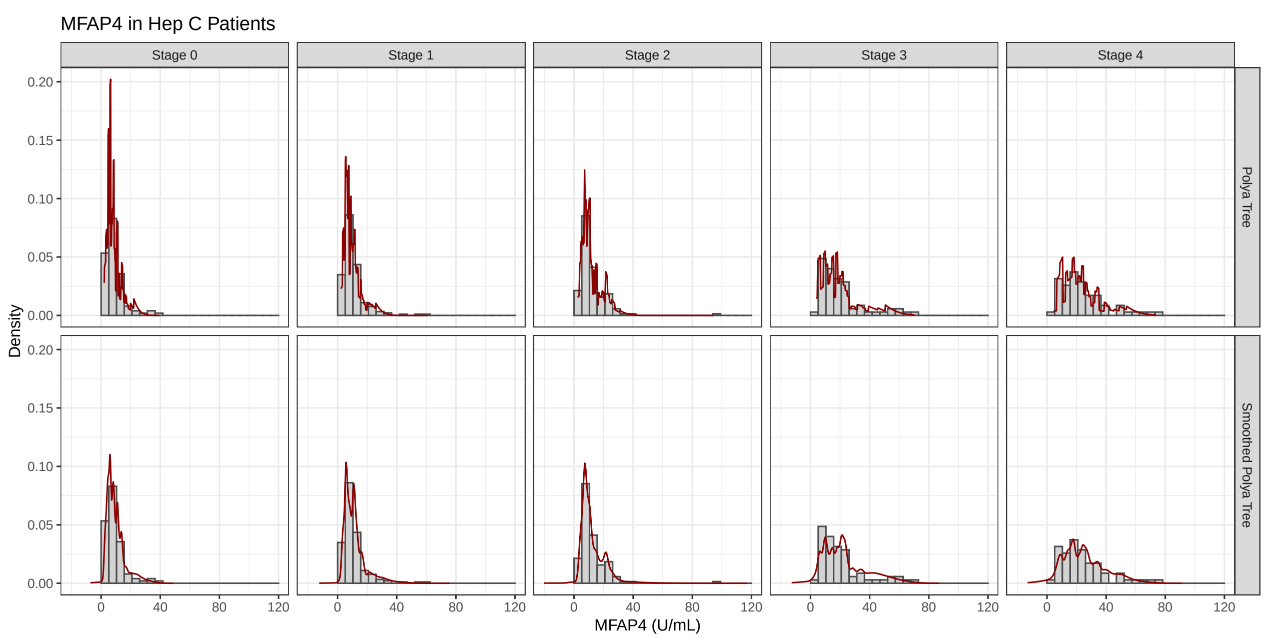

Figure 3: MFAP4 levels in $n = 542$ hepatitis-C patients by fibrosis stage with Polya tree distribution estimates (top), and smoothed Polya tree distribution estimates (bottom).

The distribution estimates themselves can provide important information about the data at hand. For example, we can conceptualize the relationship between MFAP4 levels and fibrosis severity (Figure 4). Though there is some overlap in MFAP4 levels, there is a clear difference between stages 0-2 and 3-4, particularly among the smoothed estimates. While this pattern is evident when assuming a normal distribution, the distributions are inaccurate and recognizes only two treatment groups. The versatility of these methodologies makes them relevant to several applications I outline in the article, i.e., modeling regression error, fitting random intercept generalized linear mixed models (GLMMs), and meta-analyses.

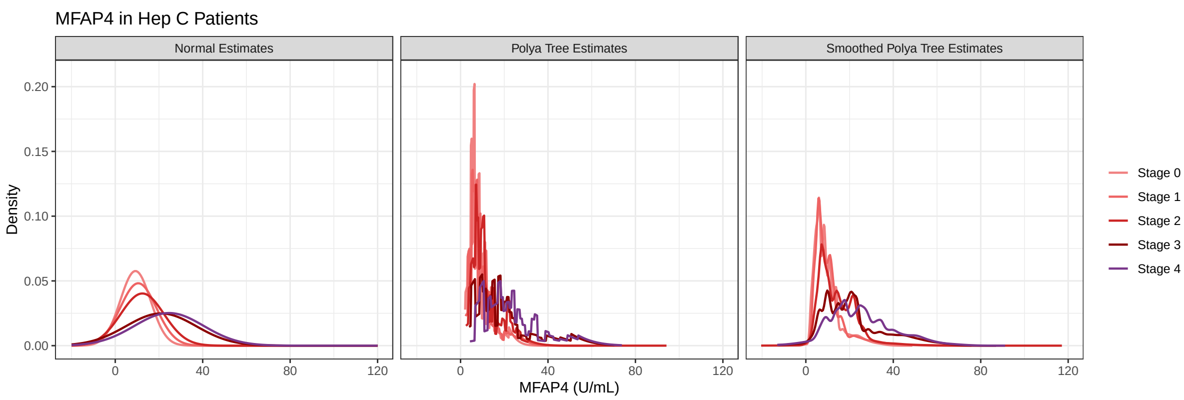

Figure 4: MFAP4 levels in $n = 542$ hepatitis-C patients by fibrosis stage (left), with Polya tree distribution estimates (middle), and smoothed Polya tree distribution estimates (right).

While there is a rich history of methodological development for univariate Polya trees, comparatively little research has been done in the multivariate setting. Multivariate Polya trees are similar in theory to univariate Polya trees, but the computational requirements are more significant, and the models are more complicated. Conceptually, understanding a multivariate Polya tree begins with a multivariate histogram. For example, we might consider other aspects of the hepatitis C patients in the MFAP4 data, e.g., age, sex, hepatitis genotype, or presence of hepatocellular carcinoma. This complicates our analysis as our picture gains a dimension for each added variable.

In Cipolli and Hanson (2019), I extend multivariate Quadratic Discriminant Analysis (QDA) (Cox, 1958) to the Bayesian nonparametric setting. In this approach, I smooth the multivariate Polya tree distribution estimate by averaging distribution estimates from two rotations, i.e., the orientation of the standard bivariate histograms in Figure 5 (left) and another on a grid that is rotated 45 degrees. The proposed method is fast and straightforward to implement, making it one of the most user-friendly approaches to supervised learning.

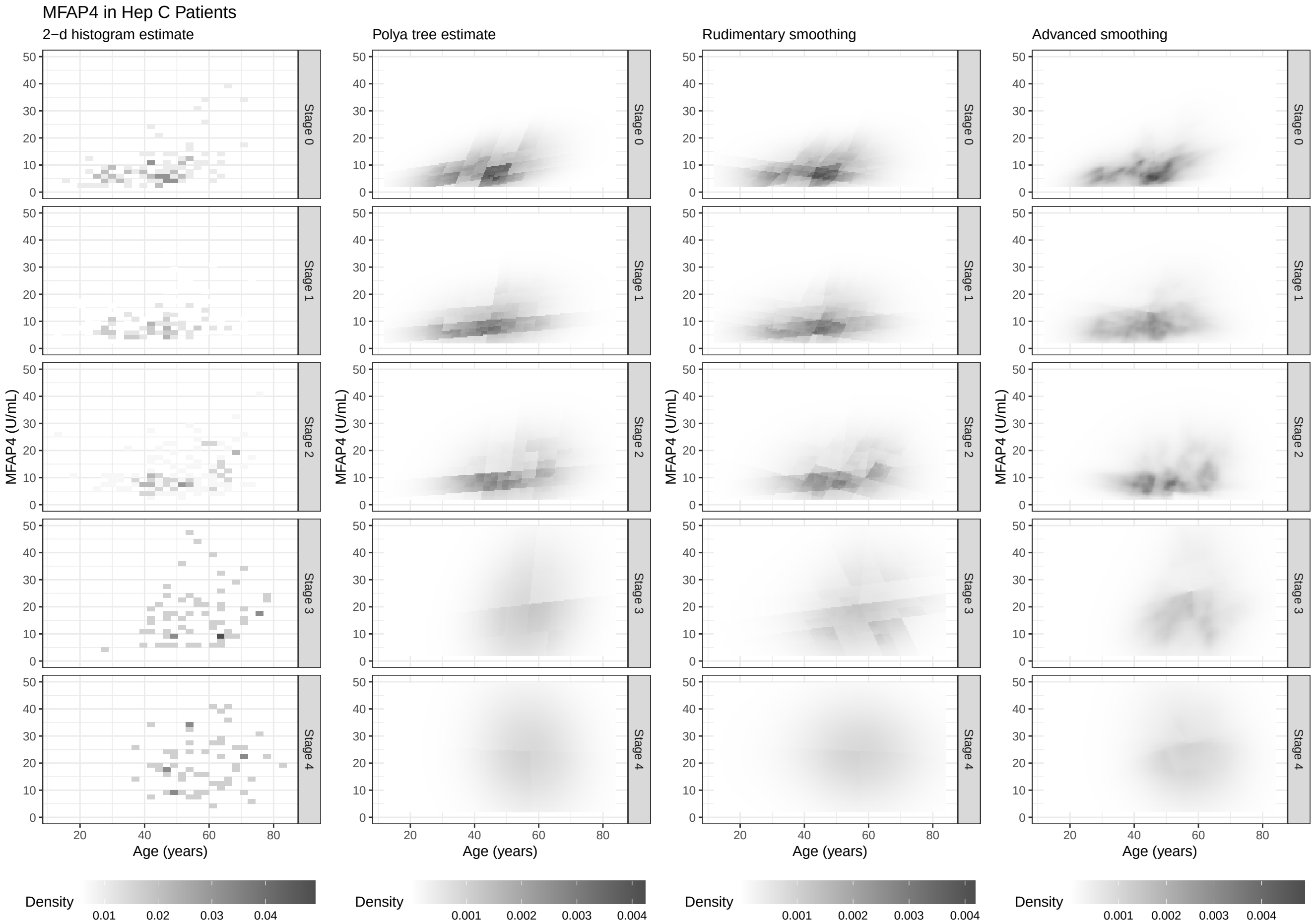

Figure 5: MFAP4 levels and age of $n = 542$ hepatitis-C patients by fibrosis stage with a standard Polya tree estimate (middle-left), with rudimentary smoothing (middle-right), and advanced smoothing (right).

Similar to the univariate setting, these multivariate estimates are spiky, which is seen in the visual artifact of the grids in the Polya Tree estimate. The solution to smoothing in the univariate case, however, is not tenable in the multivariate setting. Instead, another solution for smoothing that is justified probabilistically has to be developed. While frameworks for smoothing exist, they are problematically slow, require substantial computation time, and don't yield noteworthy smoothing.

Like the univariate setting, the smoothness makes the analysis less noisy, and the computational savings lower the barrier for use. For example, we may want to include age in our analysis of the fibrosis disease in hepatitis C patients because age and MFAP4 levels are related (Figure 5). In my recent methodological article submitted to *Statistics and Computing* (2021), I develop such a framework and propose a novel sampling technique for improving computation time from days to minutes; the smoothing results are demonstrated in Figure 5 (right). This approach attains accuracy similar to state-of-the-art prediction models without assumptions about the distribution of the data.

This inventive solution to smoothing multivariate Polya-tree-based distribution estimates in a computationally efficient way provides a rich ground for further research, as distribution estimation is key to many problems in statistics. I am currently working on two such developments. I am working with Ciaran Evans (Wake Forest University) on a framework for detecting distribution changes in real-time. This framework can be used for detecting a change in electroencephalography (EEG) output when monitoring a patient's brain function, i.e., detecting an epileptic seizure. The second project involves joint work with Zichen Ma (Clemson University) to develop an independence test using the smoothed Polya trees estimation techniques I've created. In addition to traditional independence tests, we aim to assess conditional dependence – a relationship between data dependent on another piece of data. Such procedures are essential for research in the social sciences where effects of certain conditions can vary by group; e.g., lower subjective socioeconomic status within one's racial group leads to poorer health outcomes in White, but not Black, Americans.[1]

Additional future work involving Polya trees includes extending the result for smoothing univariate Polya tree distribution estimates to the heteroskedastic case for flexibly modeling survival data, and constructing multivariate frameworks for (1) variable selection, selecting which variables to include in a model; (2) clustering, grouping observations without any prior label information; and (3) mediation and moderated mediation techniques for assessing indirect effects.

Statistical and Mathematical Collaborations

I also collaborate on projects that use parametric statistical approaches. My broad statistical training allows me to be a valuable partner in deriving and evaluating new techniques. Simultaneously, my focus on statistical assumptions and their potential effects inform my future research.

In Robertson et al. (2021), with my colleague Aaron Robertson (Colgate University), we use robust simulation studies to support using the Delaporte distribution for approximating the distribution of monochromatic complete subgraphs and arithmetic progressions. This opens the door for further research exploring how this approximation could provide better estimates of Ramsey numbers.

Further, in Bower et al. (2018a,b), with my colleague Roy Bower (Furman University), we present two new testing procedures for the independence of bivariate Poisson data, which provides researchers a way of examining the relationship between two counts (e.g., the number of doctor visits and the number of diagnosed conditions). In both projects, I designed and coded the simulation studies that displayed the effectiveness of these models compared to existing alternatives in terms of power and significance, and I assisted with the derivations. Currently, we are running simulations to assess (1) zero inflation in bivariate Poisson data, (2) dependence in zero inflation Poisson data, and (3) troubleshooting derivations for a test on bivariate Pareto data. Further, we have discussed several projects, including expanding our results to the more flexible negative binomial and Delaporte distributions. While the inability to travel in the last year has slowed our work, we plan to meet and make significant progress on our collaborative work this fall.

Multidisciplinary Collaboration

Multidisciplinary collaboration is helpful both to me and to the collaborator(s) from other disciplines. I work as a partner with researchers to adapt their qualitative hypotheses for quantitative study, execute statistical analyses, write up the results in understandable terms, and assist in drafting the manuscripts. Through these collaborations, I also find motivation for new questions in statistics and recent interesting examples for teaching.

My presence as a collaborator also expands the questions researchers can address by introducing statistical, data-heavy approaches. The statistical work I do has proven valuable to research in the natural and social sciences. This combination of experiences has led to a rich series of quantitative psychology collaborations, which contain themes from both.

Natural Science Collaborations

I helped evaluate the accuracy of diagnostic methods for carpal tunnel syndrome with John Fowler (University of Pittsburgh) using Bayesian latent class analysis (Fowler et al., 2015). This allowed for the assessment and confirmation of Dr. Fowler's hypothesis that nerve conduction velocity tests, which are uncomfortable for patients, are no longer necessary. Ultrasound measurement provides similar results and is far less invasive.

Recent work with Tina Davlin-Pater (Xavier University) and others evaluates the lasting benefits of metacognition training for kinesiology students in training, using a multi-level beta regression. This work, accepted at the *Journal of Occupational Therapy Education*, shows that students improved regulation of cognition strategies after the instruction of general self-regulated learning skills and six months after the completion of the course.

Collaborative work with Ana Jimenez (Colgate University) employs exploratory data analysis and feature selection techniques to elucidate the potential underlying causes of the decreased lifespan of large dogs compared with small breed dogs (Jimenez et al., 2018, 2021a,b). We are currently working on a project using new data to provide a more fine-grained analysis of the complete lifespan.

I also have an article submitted to Environment and Behavior with Weston Dripps (Furman University) and others that elucidates the relationship between sustainability behaviors and the presence of homeowners' associations (HOAs). This work shows how HOAs currently inhibit sustainability and proposes changes to increase sustainability practices.

In joint work with Joe Levy (Colgate University) and others, we explored debris-covered glacier deposits on Mars (Levy et al., 2021). Using a clustering approach, we show that these deposits likely formed over punctuated episodes, like those on Earth. While previous work has modeled the last 20 million years, this article provides the first evidence about the orbit of Mars and its history of ice ages over hundreds of millions of years.

Social Science Natural Language Processing Collaborations

I have published two peer-reviewed articles collaborating with my colleague Derek Silva (King's University College at Western University). In Silva et al. (2018) we employed a classification scheme to demonstrate that "scouts," expert analysts of high school athletes, promote a system of scrutiny that is arbitrary. In Kennedy et al. (2019), we employ a quantitative approach that uses natural language processing to explore how a sporting tragedy – the bus accident involving the Humboldt Broncos junior hockey team – constructs expressions of national and regional identity through public displays of collective mourning that reinforced exclusive ideals of being "properly" Canadian.

Dr. Silva and I were awarded $55,532 from the Social Science and Humanities Research Council of Canada to document how radicalization has emerged as the dominant logic for issues related to terrorism and to elucidate how certain groups strategically communicate, mobilize, and deploy resources to frame the public understanding of radicalization. This work involves traditional qualitative studies, state-of-the-art statistical approaches to natural language processing, and social network analysis to model the contours of mainstream and political discourse.

Inspired by my work with Dr. Silva, I started conversations with my colleagues Andrew Pattison (Colgate University) and Jose Marichal (California Lutheran University) on the connection between qualitative and quantitative approaches. We have worked to provide a quantitative approach that complements the work Gupta et al. (2018) using fracking policy in New York State as a case study. This article was recently accepted at Review of Policy Research, and we have several ideas for expanding this approach. For example, our current process requires hand-coding of tweets into pro- or anti-fracking narratives. A method that employs classification techniques to code tweets programmatically would increase the amount of data without increasing hand-coding.

This work about policy changes and discourse proved helpful in joint work with Hayden Smith (University of South Carolina). Our work, in revision at Policy & Internet, explores a corpus of tweets surrounding Instagram and Facebook's commitment to banning all graphic images of self-harm on their platforms. We show that while those making decisions have decided that removing all self-harm content to be the best approach, that policy removes/censors the support channels for some who use Twitter for social support and celebrating recovery.

Joint work with Donald Sull (Massachusetts Institute of Technology) defines corporate culture using posts on an employee review website. We have created data from the free-response questions and use a Bayesian approach to model the ordinal culture score (1-5). We are currently working to submit two papers based on these results. In the first, we identify five *toxic* attributes that drive negative evaluations of culture – disrespectful, exclusionary, unethical, cutthroat, and abusive. In the second, we evaluate the differences between professional employees (e.g., engineers, managers, etc.) and front-line employees (e.g., maintenance workers, unskilled laborers, etc.).

Quantitative Psychology Collaborations

In my quantitative psychology collaborations, I provide access to new and advanced statistical techniques to fortify evidence for theory-driven hypotheses. This requires me to be part of the entire research process and provide a sounding board for the initial research questions, and I am heavily involved in crafting experimental and analysis plans.

As an example of this work, we have built a team of collaborators covering social and developmental psychology – Dr. Cooley, Lauren Philbrook (Colgate University), Jazmin Brown-Iannuzzi (University of Virginia), and Ryan Lei (Haverford College) – to investigate how *White=Rich, Black=Poor* stereotypes develop and affect race-related issues in the United States. We received a major research grant of $160,000 from the Picker Interdisciplinary Science Institute to fund multidisciplinary research exploring the psychological processes that link race with class to better understand – and thus intervene to mitigate – distress and discrimination experienced by US citizens. I play a crucial role in planning the experimental design and statistical analyses to disseminate the results critical to these projects. Further, I work to incorporate new state-of-the-art statistical approaches into our projects, e.g., brief reverse correlation (Schmitz et al., 2021).

Thus far, this work has led to three publications in top journals of psychology (Cooley et al., 2019a, 2021a,b). These projects focus on the effects of subjective socioeconomic status. We have found that White participants tend to think that most White Americans are of higher status than the self, while Black participants believe that most Black Americans are of lower status. In our first project, we showed that when White Americans feel that they are falling behind the high status of their racial group, this predicts reduced stereotyping of welfare recipients as lazy, and thus, greater welfare policy support. Our second project showed that this perception leads to changes in self-reported health outcomes, e.g., emotional, mental, and physical health. In a third project, we have collected longitudinal data to better assess whether these conclusions reflect causal processes that unfold over time.

Given that we have been working with large samples of representative data, I have played a key role in testing various models while ensuring adequate statistical power, meeting statistical assumptions and corrections for multiple tests. Furthermore, I have been integral to assessing how these processes unfold over time by applying time-series analyses to recent data we collected from White and Black Americans over the months leading up to and following the most recent presidential election. We anticipate that these rich data will yield a variety of consequential publications. This broad line of work will form the basis for external grant applications that we are now drafting for The Russell Sage Foundation and NSF.

Working with a network of psychologists has led to further collaboration outside of the Picker Grant as I meet researchers during conference presentations or through being recommended by my current collaborators. For example, I have worked with Nava Caluori (University of Virginia) and others to model the relationship between globalization and prejudice worldwide using multilevel models and moderated mediation models to better explain causal processes (Caluori et al., 2020). Further, I continue to work with Dr. Brown-Iannuzzi to evaluate how people's mental visualizations of "welfare recipients" and "illegal voters" shape their policy attitudes. To do this, we have been using reverse correlation (RC) techniques. Within these tasks, participants are asked to choose, across many different trials, which of two images looks more like a "welfare recipient," for example. The images presented to participants on each trial are generated by superimposing random visual noise over a merged photo representing people of different races and genders. Across many trials, we can then average participants' image selections to get an average image/composite of whom people envision when they think of different social categories. For example, Figure 11 depicts the average image participants envisioned when asked to consider a "social security recipient" and "not a social security recipient." Critically, we have further linked characteristics of these photos (e.g., skin tone) to participants' degree of belief that social security recipients are hardworking. Dr. Brown-Iannuzzi, Jeremy Cone (William's University), and I are currently working on a grant to further develop measures and tools for the RC procedure.

Figure 11: Averaged images of those who do and do not receive social security. Consistent with expectations, social security recipients were, on average, perceived as more hardworking and Whiter than those who do not receive social security.

Some of the conversations stemming from ongoing collaborations have led to several large quantitative projects. For example, in Cooley et al. (2019b) we modeled data from millions of police stops recorded within the NYC stop-and-frisk database to find that racial biases observed in the stops were amplified for Black civilians stopped among groups (vs. individually). Additionally, in an article with Keith Payne (University of North Carolina) and others, which is in revision at *Journal of Experimental Psychology: General*, we provide insight into a pressing societal question: how is racial discrimination in the United States distributed? That is, do a few bad apples perpetuate discrimination? Or more widespread? To assess this complex research question, we have conducted meta-analyses and robust simulation studies. While some researchers have argued that discrimination is only enacted by only a few discriminatory individuals (via the Pareto principle), we find evidence that the distribution of discrimination is more aligned with a normal distribution, reflecting more widespread discrimination (see Figure 12). Our team has answered consequential research questions about the nature of contemporary racial discrimination using available data on hiring and housing discrimination through this statistical work.

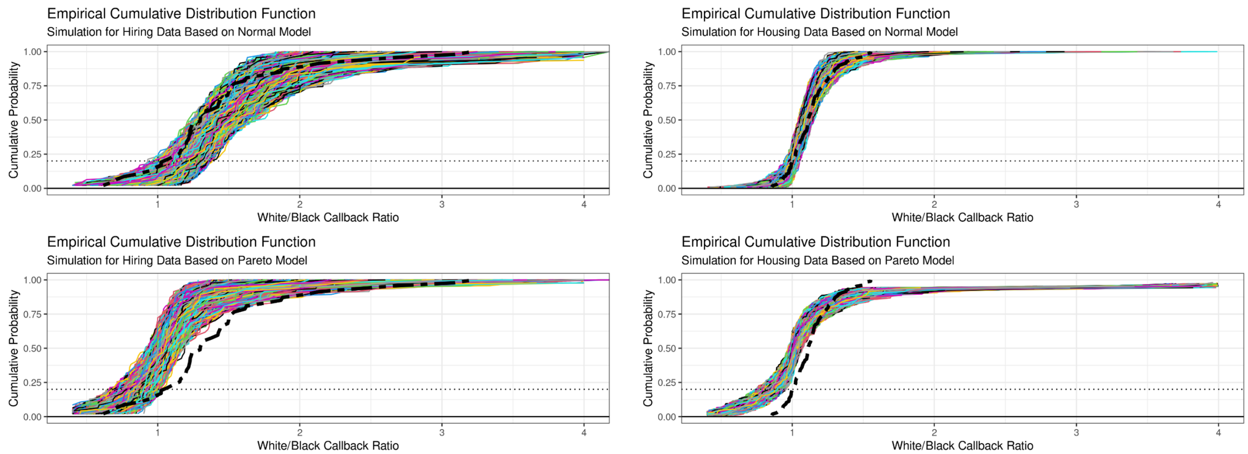

Figure 12: The empirical cumulative distribution of the discrimination observed in data is plotted using a black two-dash line and the simulations are plotted in color. Simulations under the normal model (dispersed discrimination) are more consistent with the observed data than the Pareto model (condensed discrimination).

The Data Science Collaboratory at Colgate University

Building upon the intent of my collaborative work, I am also working to develop resources that will help researchers who regularly work with quantitative data improve their data-driven processes. In pursuit of this goal, Josh Finnell (Colgate University) and I have been awarded a new initiatives grant for $3,000 from the Central New York Library Resources Council (CLRC) titled, "The Data Analysis and Collaboration Network." We have built an inter-institutional statistics collaboratory of "isolated" statisticians and data librarians from community colleges and regional universities to provide resources and collaborative support for researchers.

We have also worked with students to develop applications made freely available on the soon-to-launch resource page. These applications will make complicated discipline-specific statistical methodology more accessible using `R Shiny` to build interactive web applications (Figure 13). We hope this will be the start of an accessible resource that provides standard and advanced statistical procedures for natural and social scientists and also humanists, making Colgate a place where data science is accessible to everyone.

We have used this pilot project as the foundation of a Spencer large research grant on education to help solidify, expand, and sustain the cohort of experts to develop the next generation of data scientists.

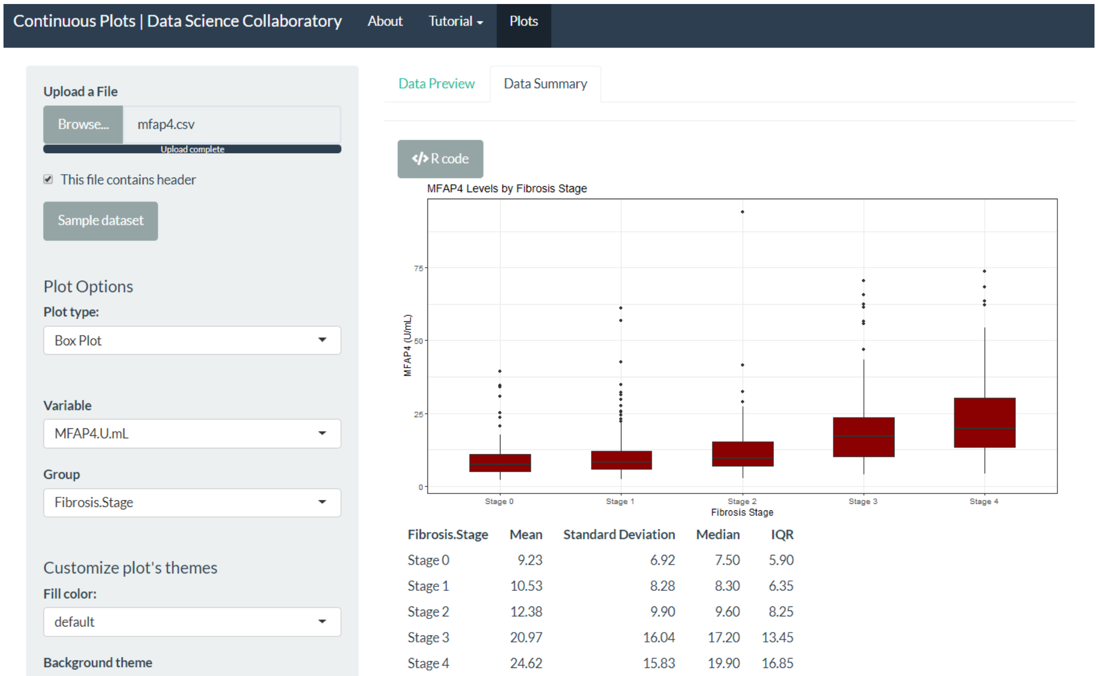

Figure 13: An example collaboratory app. This application plots data, and currently shows a boxplot of the MFAP4 data as in Figure 1.

Statistics Education

I also make a major contribution by creating course notes that I make available to students and faculty teaching similar courses. I have created notes for MATH 105 (> 500 pages), the MATH 316/416 sequence (> 800 pages), and MATH 354 (> 700 pages). My course notes are textbook-quality summaries of examples, ideas, and theories written using LaTeX. All three sets of lecture notes are close to submission-ready, and I plan to submit them as textbooks post-tenure.

Service

Presentations at Professional Meetings or Conferences

I have given twenty-one research presentations – fourteen contributed talks and seven posters at professional meetings by consistently giving presentations at our profession's largest meetings: The Joint Statistical Meetings and the Joint Mathematical Meetings and specialized conferences such as Bayesian Nonparametrics and Objective Bayes. Further, I recently was a panelist for an expert roundtable on the promises and problems of data science in the service of social justice during the keynote session at UPSTAT 2021.

I have also helped plan and organize sessions at major conferences. As a participant of Project NExT, I planned a session at the 2019 Joint Mathematics Meetings entitled "Standards-based Grading: Tools and Tips for Successful Implementation." Further, at the 2020 Joint Statistical Meetings, I organized and participated in a panel entitled "Balderdash, Codswallop, and Malarkey" for discussing methods for educating our students to use statistical reasoning and critical thinking beyond the professional realm as global citizens and decision-makers.

Editorial Service

I serve as an associate editor for the *Journal of Statistics and Data Science Education* and *The American Statistician*. I am also an active referee for articles and textbooks, as outlined in my CV. In addition to this traditional service as a reviewer, I have participated as a Scientific Review Officer on a panel for "The Small Business Disease Prevention and Management, Risk Reduction and Health Behavior Change" grants through the National Institutes of Health (NIH) in 2017 and 2020. I remain an active reviewer available for future panels.

Service at Colgate

I have actively served on a variety of internal committees. I was appointed to a three-year position on the Faculty Development Council (2017-2020), an ongoing role on the Research Computing Committee (2017-Present), and as an IRB member (2022-present). I have been elected for a three-year position on the Assessment Committee (2020-2023) and as a stand-in member of the Committee on Information Technology (2021).

Further, I have actively participated in the Colgate community by hosting Commons events and participating in many recruitment endeavors for faculty and students. I also regularly attend White Eagle and teaching tables hosted by the Center for Learning, Teaching, and Research (CLTR). Further, I share my insights with current faculty by hosting events for the CLTR and new faculty during orientation events.

Finally, I am dedicated to working together in pursuit of departmental goals. I am quick to serve the department and support my colleagues (e.g., representing the department during April Visit Days). I have served as the minute taker at meetings and as the MAA liaison for the department. I have also served on two search committees and will do so again in the upcoming academic year. Likewise, I mentor a steady stream of students interested in data science, and I have led student trips to conferences and competitions, e.g., the DataFest17 competition.

Conclusions

As this statement delineates, I have been very active as a teacher, scholar, and community member. I work to think critically and constantly about improving students' experiences in my classrooms, and my research spans purely statistical topics to myriad multidisciplinary collaborations. Moreover, our work on the Collaboratory unites my passion for teaching and research to advance the statistical integrity of the many disciplines that use quantitative data. In all these pursuits, I look forward to continuing to grow within the supportive environment of Colgate University.

Bibliography

Bower, R., Hussey, J. R., Zhang, J., Quattro, J., Cipolli, W., and Hardin, J. W. (2018a). A copula approach for testing independence between Poisson distributed variables. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 4(1):9–17.

Bower, R., Hussey, J. R., Zhang, J., Quattro, J., Muhling, B., Cipolli, W., and Hardin, J. W. (2018b). The score test for independence of two marginal Poisson variables. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 4(3-4):164–172.

Caluori, N., Brown-Iannuzzi, J. L., and Cipolli, W. (2020). Economic inequality shapes the relationship between globalization and prejudice. *Social Psychological and Personality Science*, pages 1–10. First published online.

Cipolli, W. and Hanson, T. (2017). Computationally tractable approximate and smoothed Polya trees. *Statistics and Computing*, 27(1):39–51.

Cipolli, W. and Hanson, T. (2019). Supervised learning via smoothed Polya trees. *Advances in Data Analysis and Classification*, 13(222):877–904.

Cipolli, W., Hanson, T., and McLain, A. C. (2016). Bayesian nonparametric multiple testing. *Computational Statistics & Data Analysis*, 101(1):64–79.

Cooley, E., Brown-Iannuzzi, J. L., Lei, R. F., and Cipolli, W. (2019a). Complex intersections of race and class: Among social liberals, learning about White Privilege reduces sympathy, increases blame, and decreases external attributions for White people struggling with poverty. *Journal of Experimental Psychology: General*, 148(12):761–769.

Cooley, E., Brown-Iannuzzi, J. L., Lei, R. F., Cipolli, W., and Philbrook, L. E. (2021a). The policy implications of feeling relatively low versus high status within a privileged group. *Journal of Experimental Psychology: General*, pages 1–17. First published online.

Cooley, E., Brown-Iannuzzi, J. L., Lei, R. F., Philbrook, L. E., Cipolli, W., and McKee, S. E. (2021b). Investigating the health consequences for White Americans who believe White Americans are wealthy. *Social Psychological and Personality Science*, 12(3):371–382.

Cooley, E., Hester, N., Cipolli, W., Rivera, L. I., Abrams, K., Pagan, J., Sommers, S. R., and Payne, K. (2019b). Racial biases in officers' decisions to frisk are amplified for Black people stopped among groups leading to similar biases in searches, arrests, and use of force. *Social Psychological and Personality Science*, 11(6):761–769.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20:215–242.

Fowler, J. R., Cipolli, W., and Hanson, T. (2015). A comparison of three diagnostic tests for carpal tunnel syndrome using latent class analysis. *JBJS*, 97(23):1958–1961.

Gupta, K., Ripberger, J., and Wehde, W. (2018). Advocacy group messaging on social media: Using the narrative policy framework to study Twitter messages about nuclear energy policy in the United States. *Policy Studies Journal*, 46(1):119–136.

Jimenez, A. G., Downs, C. J., Lalwani, S., and Cipolli, W. (2021a). Cellular metabolism and IL-6 concentrations during stimulated inflammation in small and large dog breeds' primary fibroblasts cells, as they age. *Journal of Experimental Biology*, pages 1–8. First published online.

Jimenez, A. G., Lalwani, S., and Cipolli, W. (2021b). Effects of metformin, rapamycin, and resveratrol on cellular metabolism of canine primary fibroblast cells isolated from large and small breeds as they age. *GeroScience*, 59(1):1–14.

Jimenez, A. G., Winward, J., Beattie, U., and Cipolli, W. (2018). Cellular metabolism and oxidative stress as a possible determinant for longevity in small breed and large breed dogs. *PloS one*, 13(4):1–20.

Kennedy, L., Silva, D., Coelho, M., and Cipolli, W. (2019). We are all broncos: Hockey, tragedy, and the formation of Canadian identity. *Sociology of Sport Journal*, 36(3):189–202.

Levy, J. S., Fassett, C. I., Holt, J. W., Parsons, R., Cipolli, W., Goudge, T. A., Tebolt, M., Kuentz, L., Johnson, J., Ishraque, F., Cvijanovich, B., and Armstrong, I. (2021). Surface boulder banding indicates Martian debris-covered glaciers formed over multiple glaciations. *Proceedings of the National Academy of Sciences*, 118(4):1–7.

Robertson, A., Cipolli, W., and Dascălu, M. (2021). On the distribution of monochromatic complete subgraphs and arithmetic progressions. *Experimental Mathematics*, 30(1):135–145.

Schmitz, M., Rougier, M., and Yzerbyt, V. (2021). Introducing the brief reverse correlation.

Scott, J. G. and Berger, J. O. (2006). An exploration of aspects of Bayesian multiple testing. *Journal of Statistical Planning and Inference*, 136(07):2144–2162.

Silva, D. M., Bower, R., and Cipolli, W. (2018). In search of a five-star: The centrality of body discourses in the scouting of high school football players. *Sociology of Sport Journal*, 35(4):324–333.