Application for Tenure-Track Instructor Position in Statistics at UBC

Vincenzo Coia

2019-12-03

Contents

Chapter 1

Cover Letter

NOTE: THIS APPLICATION IS A WORK-IN-PROGRESS

Dear members of the search committee,

I am writing to enthusiastically apply for a tenure-track instructor position in Statistics (Job #35876). Ever since joining the department of Statistics full time in early 2017 to advance our Data Science initiatives, I've come to realize just how ideal of a fit the teaching stream is for me. My aim with this application is to indicate why this is so, as well as why the department and UBC (and external to UBC!) would benefit by promoting me to the teaching stream.

Ultimately, in this position, I envision advancing all branches of Data Science education at UBC, especially through Statistics for Data Science. I'm equally excited to develop the new Minor in Data Science program as I am to continue developing the Master of Data Science (MDS) program. In fact, I'm actually rather torn between the two: there is still much to develop with MDS, and I'm rather invested in the program; but the allure of developing a new Minor program is a new equally exciting challenge that I can make valuable contributions to. Regardless of my level of involvement in these, I intend to keep an open stream of communication between the two programs, along with STAT 545A/547M, so that data science education at UBC can overall grow as a cohesive whole, as opposed to competing pieces.

My vision for doing this

I've identified four pillars of my career, which I hope will lay the foundation for my vision going forward in the teaching stream, as well as convey fit on my end. They are, in order of culmination:

1. **Application**: addressing real problems with data science.

This means continuing to work with external partners, such as the Master of Data Science (MDS) capstone partners, to understand their pressing data issues,

and lend a hand once in a while. This is important for keeping myself and data science education at UBC relevant, as well as continually expanding my skillset. Besides having mentored MDS capstone projects for the past three years, I also did some work on flood forecasting with BGC Engineering this past summer.

2. **Research**: discovering better ways to do data science.

I've come to learn that there is a deficit in figuring out how to apply Statistics to a real problem. Statistical Science (importantly) is generally about setting up a framework of stochastic assumptions, and describing properties and methods under this framework. An example of a framework is linear regression; another, a Markov model. But the reality is that real examples don't neatly satisfy assumptions. Instead, some assumptions are better *approximations* than others, yet others have different *uses* than others. Thinking about assumptions in this respect brings about questions and ways of thinking that have traditionally not been adequately addressed in the literature nor the classroom, yet are defining for Statistics for Data Science – hence why I think of this as "research". I call this approach a "problem-first" approach, as opposed to a "method-first" approach taken by Statistical Science. Even Applied Statistics may seem to take a "problem-first" approach, but the motivating real-world problem tends to be used as motivation for setting up a framework of assumptions, for which properties are then described.

I've been developing this way of thinking with MDS, but am nowhere near complete. I wish to continue this approach throughout my career.

On a smaller scale, I find that distributional thinking needs improvement. To this extent, I am developing an R package that works with distributions as objects, as opposed to handling the unwieldly pieces of functions like `dnorm()` and `rnbinom()`. My goal is to submit this package to CRAN and ROpenSci, and perhaps grab the attention of RStudio and the `tidymodels` team. Furthermore, I believe that copulas are currently not accessible enough due to complicated formulas and jargon, yet are tremendously useful for distributional forecasting. I hope to make this framework more accessible by developing an R package called `coperate` and `cmc` for easy manipulation of copulas, and easily created copula-based models (respectively). Again, submitted to CRAN and ROpenSci.

3. **Development**: turning new research into programs and tools, which are ideally accessible to the public.
4. **Teaching**: spreading data literacy and enthusiasm to the public.

- excellence in, and commitment to, undergraduate teaching;
- ability to work collaboratively as well as independently, and to supervise the work of others;
- promise of educational leadership at UBC and nationally or internationally;
- potential for creating innovative instructional environments.

Lastly, I'd like to elaborate on how my skills will be an asset to the department.

Excitement and promise for starting up the minor program, helping MDS and STAT 545A/547M evolve along with it.

hyperlinks to one or two examples of my teaching materials. Note that I don't see any of these as being "finished", but rather a work-in-progress.

- For STAT 545A/547M (Exploratory Data Analysis), I made the lecture notes and course website. Less relevant are Lectures 11+, taught by Dr. Firas Moosvi, and Lecture 6, a guest lecture taught by my TA, Victor Yuan.
- For DSCI 551 (Probability for Data Science), I wrote the lecture notes by using Dr. Mike Gelbart's notes from the previous year as scaffolding.
- I'm writing an open-source book on regression analysis for data science, called Interpreting Regression. It's in its early stages.

What I'm comfortable teaching (useful in parallel to the minor program in DS)

Tools; my comfort level with CS. Although I don't have formal training in CS outside of perhaps a few undergraduate courses, my approach to learning these things is to identify tools and techniques that would be useful, and commit to building these tools. I prefer this approach over prescribed training such as online courses, because I find it more genuine as I encounter concepts as they become relevant. For example, I've been learning web hosting as I've been learning about Hugo, blogdown, and netlify to create websites like my homepage, course websites like STAT 545A/547M, and hosting "books" such as Interpreting Regression online. I've learned shell scripting back in my PhD, and continue to learn more, as I embrace GNU Make and workflow automation in general.

Teaching tools and methods

the names of three references who have been asked to send reference letters:

1. Dr. Tiffany Timbers (current supervisor)
2. Dr. Michael Gelbart (current supervisor)
3. Dr. Harry Joe (PhD supervisor)

I am eager to begin the position on the indicated July 1, 2020 start date, but am ultimately flexible.

Enthusiastically yours,

Dr. Vincenzo Coia

Chapter 2

Curriculum Vitae

2.1 Work Experience

2017/01 - present
Lecturer of Data Science
(Initially: Postdoctoral Teaching and Learning Fellow)
Masters of Data Science Program, and the Department of Statistics
The University of British Columbia
Vancouver, BC

2009/05 - 2014/05
Short-term statistical consulting (6 projects)
UBC and private

2.2 Education

PhD in Statistics
2012/09 - 2017/02
The University of British Columbia
Conferred May 29, 2017

MSc in Mathematics and Statistics (Statistics)
2011/09 - 2012/08
Brock University
Conferred on October 13, 2012

BSc (3-year) in Biological Sciences
Minor in Earth Sciences
2005/09 - 2011/04

Brock University
Conferred "With Distinction" on October 22, 2011

BSc (Honours) Mathematics Integrated with Computers and Applications
Concentration in Statistics
2005/09 - 2011/04
Brock University
Conferred "With First-Class Standing" on June 7, 2011

2.2.1 Volunteer Positions

2016/09 - 2016/02
Science World at TELUS World of Science
Vancouver, BC
78.15 hours

2013/10 - 2014/05
Beaty Biodiversity Museum: Events Volunteer
Vancouver, BC
35.0 hours

2013/04 - 2013/09
UBC Farm
Vancouver, BC
102.5 hours

2011/06 - 2011/08
Project S.H.A.R.E. community garden
Niagara Falls, ON
15.0 hours

2.2.2 Research Assistantships

2013/05 - 2013/08
Robust penalized regression
Supervisor: Dr. Gabriela Cohen-Freue
Department of Statistics
The University of British Columbia
Vancouver, BC

2012/05 - 2012/08
2011/05 - 2011/08
2010/05 - 2010/08
Extreme value modelling
Supervisor: Dr. Mei Ling Huang

Department of Mathematics
Brock University
St. Catharines, ON

2010/09 - 2011/06
Quantum monte carlo simulations
Supervisors: Dr. Stuart Rothstein; Dr. Wai Kong (John) Yuen
Department of Chemistry and Department of Mathematics
Brock University
St. Catharines, ON

2.2.3 Teaching Assistantships

Duration: From the latter part of my undergrad, to the end of my PhD.

UBC:

- **SCIE 300: Communicating Science** (5x)

Brock University:

- **MATH 4P82/5P82: Non-parametric Statistics**
- **MATH 3P82: Regression Analysis**
- **MATH 4P81/5P81: Sampling Theory**
- **MATH 3P81: Experimental Design** (2x)
- **MATH 2F40: Mathematics Integrated w/ Computers and Applications II**

2.3 Publications and Talks

2.3.1 Articles Submitted to Refereed Journals

- Huang, M.L., Coia, V., and Brill, P.H. (2013) A cluster truncated Pareto distribution and its applications. ISRN Probability and Statistics 2013: Article ID 265373.

- Ayad, M., Coia, V., and Kihel, O. (2014) The number of relatively prime subsets of a finite union of sets of consecutive integers. Journal of Integer Sequences 17: Article 14.3.7

- Coia, V., and Huang, M.L. (2014) A sieve model for extreme values. Journal of Statistical Computation and Simulation. 84(8):16921710.

2.3.2   Articles Submitted to Conference Proceedings

- Huang, M. L., Coia, V., and Brill, P.H., A mixture truncated Pareto distribution, In JSM Proceedings 2012, Statistical Computing Section, Alexandria, VA: American Statistical Association, pp.
24882498.

2.3.3   Conference and Roundtable Contributions

- Coia, V., Nolde, N., and Joe, H. Forecasting Extremes for Flooding (Invited Talk). The 44th Annual Meeting of the Statistical Society of Canada. May 29June 1, 2016 at Brock University, St. Catharines, ON.

- Coia, V., and Jeanniard du Dot, T. (Invited Demonstration) "Using the Grammar of Graphics and Interactivity to explore Biologging Data in R". May 6, 2015. Building a Bioanalytical Theory for Analysis of Marine Mammal Movements: A Peter Wall International Research Roundtable. The University of British Columbia, Vancouver, BC.

- Coia, V. "Flood Warning: An Application of High-Quantile Regression" (Contributed Talk). SFU/UBC Joint Graduate Student Seminar (Winter). February 28, 2015 at the SFU Harbour Centre, Vancouver, BC.

- Coia, V. "A New Sieve Model for Extreme Values" (Contributed Talk). SFU/UBC Joint Graduate Student Seminar (Fall). September 29, 2012 at the SFU Harbour Centre, Vancouver, BC.

- Coia, V., and Huang, M.L. "On Estimation of Heavy Tailed Distributions" (Contributed Talk). The 40th Annual Meeting of the Statistical Society of Canada. June 36, 2012 at the University of Guelph, Guelph, ON.

- Huang, M.L., Coia, V., and Brill, P.H. "A Mixture Truncated Pareto Distribution" (Contributed Talk). The 2012 Joint Statistical Meetings. July 28August 2, 2012 at San Diego, California

2.4   Awards

2.4.1   University Issued

- 2012/09 - 2016/08: Four-Year Fellowship
- 2012/09 - 2016/08: Faculty of Science Graduate Award
- 2011/09: Dean of Graduate Studies Excellence Scholarship
- 2011/06/07: Dean's Gold Medal
- 2011/06/07: Distinguished Undergraduate Student Award in Mathematics
- 2011/03: President's Surgite Award

2.4.2 Nationally Recognized

- 2013/06: Governor General of Canada's Gold Medal
- 2012/09 - 2015/08: NSERC Postgraduate Award (Doctoral, 3-year)
- 2011/09 - 2012/08: NSERC Alexander Graham Bell Canada Graduate Scholarship (Masters)
- 2010/05 - 2010/08: NSERC Undergraduate Student Research Award

2.4.3 Individual Donors

- 2012/04: Dr. Jack Lightstone & Dorothy Markiewicz Scholarship
- 2012/04: Dr. Raymond & Mrs. Sachi Moriyama Grad. Fellowship
- 2012/03: Tomlinson Entrance Scholarship for Excellence in Mathematics and Science
- 2011/06: John and Roslyn Reed Book Prize
- 2010/09: Art Bicknell Scholarship in Mathematics
- 2010/09: Ian D. Beddis Family Scholarship
- 2010/09: Terry and Sue White Mathematics and Science Scholarship
- 2007/09: M.J. ("Mel") Farquharson Scholarship
- 2006/09: Scholler Foundation Scholarship in Chemistry

2.4.4 Athletic Awards

- 2016/04/03: Sportsmanship Award, UBC Thunderbirds Sport Clubs (Fencing)
- 2016/02/28: Bronze Medal in Senior Mixed Epee (Intercollegiate Tournament, UBC)
- 2015/11/16: Bronze Medal in Senior Mixed Epee (Remembrance Day Tournament, UBC)
- 2012/03/28: RM Davis Surgite Award
- 2009/12/12: First Place Award Winter Epeedemic, Toronto Fencing Club
- 2009/03: Varsity Fencing Rookie of the Year Award

2.4.5 Declined Awards

- 2012/04: Ontario Graduate Scholarship (Doctoral)
- 2011/05: Ontario Graduate Scholarship (Masters)

2.5 Professional Activities

Dept. of Statistics, The University of British Columbia:

- 2016/05 - 2016/06: **Search committee member (for CRC 2 faculty position)**
- 2015/04/30: **Organizer of the "How to Write an Awesome Abstract" workshop**
- 2014/06 - 2015/05: **Organizer of the SFU/UBC Joint Seminar**
- 2014/06 - 2014/09: **Organizer of the Graduate Student Trip**
- 2014/04 - 2015/05: **Co-founder of the Graduate Writing Forums**
- 2013/06 - 2014/05: **Statistics Graduate Student Representative**

Chapter 3

Statement of vision for education in statistics and data science

- what "educational leadership" (EL) means to you and what EL you might pursue in the future. Vision for long-term career growth towards Prof of Teaching, even if it's tentative.
  - Existing examples:
    * Curriculum development of 551 throughout the years; 561 w/ Gaby and Sunny, later Tom; 531; 511?; 563; BAIT 509; 562; STAT 545A/547M.
    * Arrangement of (mostly) statistical concepts as they relate to data science on a whole-program level.
  - Promise: distplyr?
  - Promise: Promotion of distributional forecasting methods.
  - Promise: Host conference
  - teaching, mentorship and inspiration of colleagues: post-docs, Sunny, TA's for guest lecture.
  - Making resources public: 551, 545A/547M (expanded from Jenny's work).

- Service to the academic profession, to the University, and to the community:
  - BGC project (important to stay relevant with pressing issues, and to bring industry experience to the program)
  - MDS vision statement
  - Blog post?
  - ASDa workshops in 2017.

Responsible use

Chapter 4

statement of teaching and training philosophy

My teaching takes to heart what I learned from Greg Wilson, that teaching is much more about motivating students than it is knowledge transfer. Students have access to any information they want through the internet, but a classroom environment has the powerful advantage of the presence of peers and a knowledgeable instructor. I take a semi-spontaneous approach to teaching, as inspired by improv comedy – I have a rough guideline that I abide to, but am not afraid to deviate from this where necessary, responding to how the class is feeling about a topic. I use various methods to keep constant awareness of how the class is feeling, through simple methods such as asking questions, to more modern active learning strategies such as think-pair-share, which leverages the knowledge of peers.

I believe in quality over quantity, and this is especially important for a program as time-sensitive as the Master of Data Science program at UBC that I currently teach in. Here, I focus on fundamental concepts, asking the question "what must the students absolutely know by the end of this course?" To help answer this, I look to fundamental concepts *as they relate to applications*, not necessarily how they are developed in academia. I stick to these core concepts, and show students just how far they can go by exploring deeper concepts and data science methods – again, going back to motivation over knowledge transfer.

I believe that teaching is far less effective when done "in a vacuum", as opposed to collaboratively with the input and feedback from peers. I'm lucky to be involved with open and communicative colleagues who can share their input, to build world-class courses in data science. And I'm happy to be on the other end as well, providing input and feedback to my colleagues.

Chapter 5

Diversity Statement

To me, people deserve the same respect regardless of their identity. Any differential treatment is discrimination, and is problematic because it leaves people feeling disrespected, or worse, puts an impediment on one's life.

Our society has made some great progress on inclusivity (the opposite of discrimination). Racial and sexual discrimination are significantly less these days compared to several decades ago. But the battle is still not over, as is evidenced by graffitied rainbow sidewalks and defaced mosques.

Some discrimination is still rife. Gender discrimination is one such example, which is now starting to see some progress. But there are still other types of discrimination that are not yet well known. One such example that I feel is happening is discrimination against fathers.

Not only is it healthy for each individual to receive respectful and equal treatment, we have so much to gain by having a diverse group of people around us. There's strength in diversity.

The problem goes deeper and more elusive with the well-intentioned people who still discriminate unintentionally:

The problem does not necessarily lie with a discriminative response from an individual, if the response was unintentional. Despite the best of intentions, it's our environment and society that is responsible for crafting an automatic response from an individual. So, a good-intentioned individual that does a double-take after seeing me and my male partner holding hands should not beat themselves up over it – it just means that the cumulative effect of their environment over time crafted such a response.

5.1   My experience with discrimination

As a member of the LGBTQ+ community, I continue to experience discrimination. Me and my male partner holding hands in public still forms a spectacle to many, some even stopping to watch us after we've walked by. Wedding vendors still referring to a "bride" when mentioning our wedding. Being invited to a wedding under the condition that my partner and I show no affection.

Things were worse in my adolescence, where homosexuality was "discouraged" in my environment, leaving me feeling socially out of place and fearful. Luckily, very few extended family members have a problem with my identity, and the rest embrace me.

Even though I'm a cis male, I'm quite passionate about gender issues, because they are largely not being embraced by our society, and because we're all affected by it (although transgendered and non-binary people are affected on an entirely different level).

To me, any type of brainwashing is deplorable, yet gender brainwashing is ubiquitous. You won't find a pink yoga mat in the men's section.

pink

Whistler bathroom

The solution to gender discrimination does not involve abolishing the notion of gender altogether, because gender has been proven to be important to humanity. Solution is about rather identifying a spectrum for which the extremes might be called "masculine" and "feminine".

Women's bathroom at 49th ||.

Fatherhood

5.2   Contribution

To me, contributing to inclusivity is about creating an inclusive environment, especially when it comes to making course content; calling out discrimination, even when the discrimination is unintentional; and being a role model for others by being comfortable about who I am.

5.2.1   Creating an inclusive environment

Involves:

- Changing our language to be less presumptuous. This means saying things such as "pregnant people" instead of "pregnant women", referring to one's

"partner" instead of saying "wife" or "husband" (assuming a gender), not referring to someone by their race if not relevant (such as "I was talking to a Latino man the other day, …").

- Posting online content that suggest inclusivity. For example, not using data that indicates gender as binary (because it's naive and ultimately untrue, even if gender is paraded as "sex"), and not indicating female:male ratios and parading naive terms such as "gender balanced" (because that's naive too), but rather indicating percentage belonging to a minority group (such as "percent self-identifying as female").
- Removing spaces that discriminate by gender – this means a complete decoupling of bathrooms and gender (one should never have to say "I'm not allowed in this room because I'm a man/woman"). UBC needs to first take low-hanging fruit by abolishing gender from single-person bathrooms (like we have in the stats dept), then focus on root issues of multi-person bathrooms (which at first seem gender-related, but are actually not), such as an overall lack of privacy.
- Disassociating identities with career roles.
- Including a "covenant" or "code of conduct" in collaborative (and student) projects.

5.2.2 Calling out discrimination

Means pointing out when non-inclusive language or behaviour is used, whether intentionally or not. Critically, this should be done with compassion, as opposed to accusation, because (1) people may not know any better, and (2) even if they do know better, this language can accidentally slip due to many years of belonging to a less-inclusive environment. Calling out discrimination can help educate people of the first type, as well as help re-program people of the second type. Examples include:

- Someone telling someone else that they're in the "wrong bathroom".
- Example with MDS website, and bringing up issues in our academic meetings.

"Coming out" tells others that I'm proud of who I am, but also indicates sensitivity to issues of gender and sexual identity.

- Posting my preferred personal pronouns (he/him/his) and a rainbow emoji online.
- Posting a Positive Space sticker outside of my office.
- Contributing my LGBTSTEM interview.

As a result of this, I'm hoping that others can feel safer around me, and that I inspire others to have pride in who they are as well. I also hope that it brings awareness to those who are not familiar with LGBTQ+ issues – for example, the more someone sees preferred personal pronouns being specified, the more

likely they are to look up why more and more people are posting this and what this means.