Transitioning to Research

This course is designed to facilitate the transition from course work to your dissertation research. The course centers around a project which involves replicating a published methodological research article. The project serves a vehicle for learning the key skills of conducting research, including problem selection, literature review, writing research articles and developing powerful presentations. You will develop a written report and an oral presentation during the quarter.

Instructor: **Gary Chan** (http://faculty.washington.edu/kcgchan/) (kcgchan@uw.edu)

- Class meets Tuesday and Thursday, 1.30-3.20 pm
- Location: Health Sciences Building I-142

Link to useful resources (https://canvas.uw.edu/courses/1272457/pages/useful-resources) (new!!)

Selecting a paper for 572B

You can choose from this list (https://canvas.uw.edu/courses/1272457/files/55183250/download?wrap=1) or any other statistical methodology papers (subject to the criteria below). However, if you choose a paper not from the list, you need to obtain my approval. If you identify a paper that you would like to use, please email Gary (**kcgchan@uw.edu**) (mailto:kcgchan@uw.edu)) explaining your reason/s for selecting the paper including how you have determined that you will be able to manage the work involved. I will also need you to confirm that this paper is not part of your RA work or a previous class/project. Once your selection has been approved the paper will be marked as having been selected and not available to others.

Your selected paper should satisfy the following criteria:

- A methodologically innovative approach to a relevant/interesting statistical problem
- Draws on knowledge of material mastered during 570/571
- Replication of the solution presented in the paper will constitute a theoretically and/or computationally challenging exercise.
- Work including final report and presentation must be completed by the end of spring quarter.

(Note: some papers are longer than others, some has more proofs and/or simulations than others – You will need to replicate every theorem and simulations in the paper. Reports in the previous years range from 30 to 60 pages. Points will be taken off if you cannot replicate all results by the end of the quarter. **Choose your paper wisely**.)

- New to you, i.e. not the same as, or closely similar to papers you may have studied in this way for other courses/sequences or an RA
- The paper should not have been studied in previous years. A list of previous papers is available on Canvas.
- On a topic you find of interest.

Every student needs to select their own paper. Note that the paper needs to be approved by the instructor if they are not on the list, but **you need to email the instructor with your choice whether the paper is in the list or not**. At any time including before the course starts, you may email or request meeting with the instructor to discuss your choice, and have your final choice verified and noted on the paper list by the end of the first week of class (4/5/2019).

A list of papers for 2019 suggested by Biostats faculty is here (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22enrollment_type%22:null,%22anonymous_instructor_annotations%22:null,%22submission_id%22:null,%22user_id%22:100000003257692,%22attachment_id%22:55183250,%22type%22:%22canvadoc%22%7D&hmac=e0bce142b2a5ccd77f433621c6afe0f724b78f0d) (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22enrollment_type%22:null,%22anonymous_instructor_annotations%22:null,%22submission_id%22:null,%22user_id%22:100000003257692,%22attachment_id%22:55183250,%22type%22:%22canvadoc%22%7D&hmac=e0bce142b2a5ccd77f433621c6afe0f724b78f0d) .

Papers selected in previous years are here (https://canvas.uw.edu/courses/1272457/files/55183514/download?wrap=1) (https://canvas.uw.edu/courses/1272457/files/55183514/download?wrap=1) .

Example reports from previous years:

572paper-v4.pdf (https://canvas.uw.edu/courses/1272457/files/55183459/download?wrap=1)

Example 3.pdf (https://canvas.uw.edu/courses/1272457/files/55183479/download?wrap=1)

Final Report.pdf (https://canvas.uw.edu/courses/1272457/files/55183506/download?wrap=1)

Evaluation

Presentations: first 10%, second 15%, final 25%

Report: 50%

Session#	Date	Session Title
1	April 2	Introduction Murphy (1997) (https://canvas.uw.edu/courses/1272457/files/55397659/download?wrap= Hamada and Sitter (2004) (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en
2	April 4	Structure of a statistical paper Notes (https://canvas.uw.edu/courses/1272457/files/55464354/download?wrap=1) Personal Examples: Chan et al. (2018). AoS. (https://canvas.uw.edu/courses/1272457/files/55464268/downlo Chan et al. (2016) JRSSB (https://canvas.uw.edu/courses/1272457/files/55464289/down Chan and Wang (2017) JASA (https://canvas.uw.edu/courses/1272457/files/55464299/d Chan and Qin (2016) Biometrika. (https://canvas.uw.edu/courses/1272457/files/554643 Wong and Chan (2018) Biometrika (https://canvas.uw.edu/courses/1272457/files/55464 Chan and Wang (2012) Biometrics (https://canvas.uw.edu/courses/1272457/files/55464 Chan and Wang (2010) AoAS (https://canvas.uw.edu/courses/1272457/files/55464327/d Chan and Qin (2015) Biostatistics (https://canvas.uw.edu/courses/1272457/files/554643 Yee and Chan (2016) Stat Med. (https://canvas.uw.edu/courses/1272457/files/55464341/ Learning styles and research efficacy and outcome (in psychology) (https://canvas
3	April 9	Literature review Notes (https://canvas.uw.edu/courses/1272457/files/55557240/download?wrap=1)
4	April 11	Simulations Notes (https://canvas.uw.edu/courses/1272457/files/55611250/download?wrap=1) (https:// On the Assessment of Monte Carlo Error in Simulation Based Statistical Analyse Design and Analysis of Monte Carlo Experiments Attacking the Conventional Wis

5	April 16	Student presentation 1 Xiudi Li (slides (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en) Edward Zhao (slides (https://canvas.uw.edu/courses/1272457/files/55726151/download? Eric Morenz (slides (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en) Thayer Fisher (slides (https://canvas.uw.edu/courses/1272457/files/55726155/download?
6	April 18	Student presentation 1 Serge Aleshin-Guendel (slides (https://canvas.uw.edu/courses/1272457/files/55728896/d Tianyu Zhang (slides (https://canvas.uw.edu/courses/1272457/files/55728905/download? Sijia Li (slides (https://canvas.uw.edu/courses/1272457/files/55728913/download?wrap=1
7	April 23	Student presentation 1 Spencer Hansen (slides (https://canvas.uw.edu/courses/1272457/files/55728922/downloa Kun Yue (slides (https://canvas.uw.edu/courses/1272457/files/55733330/download?wrap= Si Cheng (slides (https://canvas.uw.edu/courses/1272457/files/55733335/download?wrap Yiqun Chen (slides)
8	April 25	Student presentation 1 Xu Wang (slides (https://canvas.uw.edu/courses/1272457/files/55733456/download?wrap Angela Zhang (slides (https://canvas.uw.edu/courses/1272457/files/55733452/download? Hyunju Son (slides (https://canvas.uw.edu/courses/1272457/files/55733447/download?wr
9	April 30	Understanding statistical methods Notes (https://canvas.uw.edu/courses/1272457/files/56086886/download?wrap=1)

		Example 1 (https://canvas.uw.edu/courses/1272457/files/56086970/download?wrap=1) Example 2 (https://canvas.uw.edu/courses/1272457/files/56086983/download?wrap=1)
10	May 2	Presenting statistical methods Notes (https://canvas.uw.edu/courses/1272457/files/56146130/download?wrap=1) Example 1 (manuscript) (https://canvas.uw.edu/courses/1272457/files/56146143/downlo Example 2 (manuscript) (https://canvas.uw.edu/courses/1272457/files/56146151/downlo Example 3 (presentation) (https://canvas.uw.edu/courses/1272457/files/56146173/down
11	May 7	Student presentation 2 Yiqun Chen (slides (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en (https://canvas.uw.edu/api/v1/canvadoc_session?blob=%7B%22moderated_grading_whitelist%22:null,%22enable_annotations%22:null,%22en Serge Aleshin-Guendel (slides (https://canvas.uw.edu/courses/1272457/files/56254180/d Xu Wang (slides (https://canvas.uw.edu/courses/1272457/files/56254189/download?wrap Kun Yue (slides (https://canvas.uw.edu/courses/1272457/files/56254196/download?wrap=
12	May 9	Student presentation 2 Edward Zhao Si Cheng Xiudi Li Hyunju Son
13	May 14	Student presentation 2 Tianyu Zhang Eric Morenz Spencer Hansen
14	May 16	Student presentation 2 Sijia Li

		Angela Zhang Thayer Fisher
15	May 21	Finishing a paper
16	May 23	Submission and peer review
17	May 28	Student presentation 3 Si Cheng Eric Morenz Sijia Li Tianyu Zhang
18	May 30	Student presentation 3 Yiqun Chen Thayer Fisher Hyunju Son Angela Zhang
19	June 4	Student presentation 3 Edward Zhao Kun Yue Serge Aleshin-Guendel
20	June 6	Student presentation 3 Xiudi Li Xu Wang Spencer Hansen

Current student selections (will update frequently):
Serge Aleshin-Guendel: E. Johndrow, K. Lum and D. B. Dunson. Theoretical limits of microclustering for record linkage. Biometrika.

Si Cheng: D. Rothenhausler, N. Meinshuasen, P. Buhlmann and J. Peters. Anchor regression: heterogeneous data meets causality. arXiv.

Edward Zhao: Sun W. (2012). **A statistical framework for eQTL mapping using RNA-seq data** (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3218220/pdf/nihms-307768.pdf) . *Biometrics*.

Kun Yue: Shi P, Li H. A model for paired-multinomial data and its application to analysis of data on a taxonomic tree. Biometrics.

Xiudi Li: X. Qiao, S. Guo and G. M. James. **Functional Graphical Models** (https://amstat.tandfonline.com/doi/abs/10.1080/01621459.2017.1390466) . JASA.

Thayer Fisher: M. H. Farrell, T. Liang and S. Misra. Deep Neural Networks for Estimation and Inference: Application to Causal Effects and other Semiparametric Estimands. [**arXiv** (https://arxiv.org/abs/1809.09953)]

Steven Wang: B. Mark, G. Raskutti and R. Willett. Network estimation from point process data. [**arXiv** (https://arxiv.org/pdf/1802.04838.pdf)]

Sijia Li: K Evans, B Sun, J. Robins and E. J. Tchetgen Tchetgen. Doubly robust regression analysis for Data Fushion [**arXiv** (https://arxiv.org/abs/1808.07309)]

Tianyu Zhang: Y Vardi, Z Ying and C. H. Zhang. **Two-sample tests for growth curves under dependent right censoring** (https://www.jstor.org/stable/2673694) . Biometrika.

Angela Zhang: T. W. Randolph, S. Zhao, W. Copeland, M. Hullar and A. Shojaie. **Kernel-penalized regression for analysis of microbiome data** (https://projecteuclid.org/euclid.aoas/1520564483) . AoAS.

Spencer Hansen: Josephine K Asafu-Adjei and Allan R Sampson. Covariate adjusted classification trees[**abstract** (https://www.ncbi.nlm.nih.gov/pubmed/28520903)] *Biostatistics.*

Eric Morenz: Alam S, Moodie EEM and Stephens DA. Should a propensity score model be super? The utility of ensemble procedures for causal adjustment. Statistics in Medicine.

Yiqun Chen: Jewell S and Witten DM. Exact spike train inference via ℓ0 optimization. AoAS.

(https://docs.google.com/document/d/1_CI96TCkaC43fozjeaw8FnajV9hDCS7d83y_bVmaV6s/edit?usp=sharing)