Syllabus – STOR 890
Reading Classics:
Topics in Foundations of Statistics

Spring 2018 (January 10 – April 30)
Section 001, TuTh 9:30-10:45am
Hanes 107

Instructor:	Jan Hannig	**Phone:** (919) 962-7511	
Office:	330 Hanes	**E-mail:** jan.hannig@unc.edu	
Office Hours:	M 1:00 – 2:00pm	**Course home page on**	
	and by appointment	http://www.unc.edu/~hannig/STOR890RC	
		https://sakai.unc.edu/x/uMfd1M	

Target Audience: Ph.D. students in the Department of Statistics and Operations Research. It is assumed that students have taken enough advanced statistics courses to be able to read research articles.

Required Text:

- The list of papers we will read in this class can be found below.

Optional Text:

- Kotz, S., & Johnson, N. L. (Eds.). (2012). Breakthroughs in Statistics: Foundations and basic theory. Springer Science & Business Media.

Course Objective: In this class we will discuss statistical papers that had a big influence on the field.

Assessment: Your grade will be based on class presentation, scribing, and participation. Sign up for the tasks here: https://docs.google.com/spreadsheets/d/1-jnPg9mqkKRFJjzIbAFfACuyxNiYQBJEzVsOdOyJXt0/edit?usp=sharing

Readings and Annotation. Each student is responsible for all reading material assigned. By midnight on Sundays prior to each class, everyone will post regarding this week's reading material online on forums on sakai with at least one question, a response to a question, or a comment. The task is to identify and point out a concept or an argument that you don't understand and ask a question about it, explaining what about it you don't understand. Alternatively, you can answer somebody else's question. We will incorporate these questions into the discussion during class.

Leading a Class Discussion. Each week, a team of two students will prepare a presentation aimed for generating an interactive class discussion. Presenters assume that everyone read the material, and are prepared to critically analyze it and add insight to the reading. The presentation should highlight key results/definitions/concepts from the

reading and briefly summarize it (no more than 10 min). Presenting students will also incorporate questions posted on the forums on sakai offer their answers and/or direct them to the class. We encourage presenters to include their own questions and/or general thoughts about the assigned material, think of examples that illustrate main results, trace further development of highlighted ideas in the literature, and offer further readings for those who are interested. The presentation should be planned as one would plan a discussion section, not a lecture.

Scribing. A team of students will be assigned for each meeting to write down all the questions and answers, thoughts, claims and ideas that come up during the class. The presenting team will provide their materials to the scribing team so that they can incorporate all the notes that they took and tie everything together. The scribing team has one week to draft the notes and e-mail them to the presenting team for further revision and comments. At the end, the document should summarize the ideas presented in the class, based on the assigned reading, as well as recount the discussion that followed. Note that the discussion should not be transcribed completely word-for word, but rather in the form of a summary of the main points, although citations are permitted if necessary. The final version of the scribing is due to Jan two weeks after the presentation.

Course Outline:

Week 1

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157-175.

Student. (1908). The probable error of a mean. *Biometrika*, 1-25.

Week 2

Fisher, R. A. (1922) On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London Series A.*

Fisher, R. A. (1925). Theory of Statistical Estimation. Mathematical Proceedings of the Cambridge Philosophical Society, 22(5), 700–725.

Savage, L. J. (1976). On rereading R.A. Fisher. *The Annals of Statistics*, 441-500.

Week 3

Neyman, J., & Pearson, E. S. (1933) On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, *231*, 289-337.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. In *Proc. R. Soc. Lond. A* (Vol. 160, No. 901, pp. 268-282). The Royal Society.

Week 4

Wald, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *The Annals of Mathematical Statistics*, *10*(4), 299-326.

Birnbaum, A. (1962). On the foundations of statistical inference. *Journal of the American Statistical Association*, *57*, 269–326.

Week 5

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological review*, *70*(3), 193

Dawid, A. P., Stone, M., & Zidek, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference. *Journal of the Royal Statistical Society. Series B*, 189-233.

Week 6

Godambe, V. P. (1982). Ancillarity principle and a statistical paradox. *Journal of the American Statistical Association* 77, 931-933.

Brown, L. D. (1990). An ancillarity paradox which appears in multiple linear regression. *The Annals of Statistics*, 471-493.

Week 7

Fraser, D. A. S. (1966). Structural probability and a generalization. *Biometrika*, *53*(1-2), 1-9.

Hannig, J., Iyer, H., Lai, R. C., & Lee, T. C. (2016). Generalized fiducial inference: A review and new results. *Journal of the American Statistical Association*, *111*(515), 1346-1361.

Week 8

Tukey, J. W. (1962). The future of data analysis. *The annals of mathematical statistics*, *33*(1), 1-67

Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, *35*(1), 73-101.

Week 9

James, W., & Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, 361-379.

Efron, B., & Morris, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. *Journal of the American Statistical Association*, 68(341), 117-130.

Week 10

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*(1), 1-26.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *The Annals of Statistics*, 927-953.

Week 11

Welch, B. L., & Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *Journal of the Royal Statistical Society. Series B (Methodological)*, 318-329.

Barndorff-Nielsen, O. E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio. Biometrika, 73(2), 307–322.

Week 12

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087-1092.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1), 97-109.

Geman, S., & Geman, D. (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In Readings in Computer Vision (pp. 564-584).

Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, *85*(410), 398-409.

Week 13

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, *27*(11), 1134-1142.

Breiman, Leo. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author) Statistical science 16: 199-231.

Week 14

Aizerman, A., Braverman, E. M., & Rozoner, L. I. (1964) Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, *25*, 821-837

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

There will be guest lectures at some points during the year.

Note: The instructor reserves the right to make any changes he considers academically advisable. It is your responsibility to attend classes and keep track of the proceedings.