140.653 Methods in Biostatistics III
Regression Analysis for Continuous Responses Syllabus
Third Term, 2020-2021

Virtual Format:

Synchronous lecture (same lecture offered at two times):
Tuesdays, 7:30-8:50am EST, 10:30-11:50am EST

Asynchronous lecture replacing scheduled Thursday class session: Posted Wednesday

Synchronous lab (same lab offered at two times):
Tuesdays, 3:30-4:20pm EST, 9:30-10:20pm EST

Office hour: Thursday 10:30-11:50am EST and by appointment

Tentative class schedule:

In the meeting column, S indicates a synchronous meeting, R indicates recorded lecture

The date for the recorded session is the recommended date when students should listen to the recorded lecture (Thursdays), consistent with the in-person structure of the course. Recorded lectures will be posted on Wednesday.

Meeting	Date	Topic	Readings
1: S	Jan 26	Introduction – Scientific method – Definition of regression – Statistical smoothing ("machine learning") and model-based regressions	https://www.khanacademy.org/math/linear-algebra; HTF, Ch 1-2 FEH,
2: R	Jan 28	Basic functions for building regression models: splines; indicator variables; interactions	HTF 5.1-5.3, 5.9 FEH, 2.4
3: S	Feb 2	Question, Question, Question: addressing specific scientific questions using regression models – Comparing like-to-like; control for confounding – Effect modification – Direct and indirect effects	FEH 2.7
4: R	Feb 4	Continuation of Question, Question, Question	FEH 2.7
5: S	Feb 9	Introduction to the classical linear regression model – Specification of model – Simple linear regression – Geometry with two predictors – Least squares equations – Added variables approach – Maximum likelihood with Gaussian errors	HTF, 3.1-3.2
6: R	Feb 11	Vector representation of linear regression	HTF, 3.1-3.2

		– Notation and model specification – Least squares equations – Geometry of least squares – Distributions of key statistics	
7: S	Feb 16	Continuation of Vector representation of linear regression	HTF, 3.1-3.2
8: R	Feb 18	Advanced inference for Linear Models – ANOVA for regression models – Linear contrasts – Non-linear functions of parameters	
9: S	Feb 23	Model checking and key extensions – Mean model: residuals –vs- predicted variables: added variable plots – Variance model: residuals^2 –vs- predicted variables – Independence: autocorrelation function; generalized least squares – Influential observations: DBETAS	FEH Ch 3,7
10: R	Feb 25	Continuation of Model checking and key extensions	FEH Ch 3,7
11: S	Mar 2	Continuation of Model checking and key extensions	FEH Ch 3,7
12: R	Mar 4	Regression models for correlated responses	
13: S	Mar 9	Continuation of Regression models for correlated responses	
14: R	Mar 11	Missing data for linear regression	
15: S	Mar 16	Continuation of Missing data for linear regression	
16: R	Mar 18	Course review	FEH 4.5

Tentative Lab Schedule:

Lab	Date	Work Pending	Data Analytic Skills Covered
1	Jan 26	Prob Set 1	Guidelines for statistical reports Graphics principles; ggplot2
2	Feb 2	Prob Set 1	Cross-validation; computing CV errors
3	Feb 9	Prob Set 1	Adjusted variable plots
4	Feb 16	Prob Set 2	Sampling distributions for functions of regression coefficients; bootstrapping in R
5	Feb 23	Prob Set 2	Displays for longitudinal data; wide vs long data format; autocorrelation function
6	Mar 2	Prob Set 3	Fitting longitudinal models in R
7	Mar 9	Prob Set 3	Prob Set 3 discussion
8	Mar 16	Prob Set 4-Data analysis project	Prob Set 4 discussion

Key Due Dates:
Math Certification Quiz: By February 5th you should complete this quiz
Problem Set 1: February 11th
On-line Quiz 1: February 15th
Problem Set 2: February 25th
On-line Quiz 2: March 1st
Problem Set 3: March 11th
In-class Quiz 3: March 12th
Problem Set 4: March 19th

Office Hours

Elizabeth Colantuoni: Thursday, 10:30-11:50 AM EST and by request

Books for Reference

(HTF): Hastie, T, Tibshirani, R, Friedman, J. 2013. The Elements of Statistical Learning. Springer.http://statweb.stanford.edu/~tibs/ElemStatLearn/index.html

- Provided pdf copy in on-line library

(FEH): Harrell, FE Jr. (2015). Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer. http://link.springer.com/book/10.1007%2F978-3-319-19425-7

The Statistical Sleuth: A course in methods of data analysis, 3rd edition. 2013 Fred Ramsey and Dan Schafer. Brooks/Cole Cengage Learning.