

Violations of model assumptions

or "just" improving models

Suspicion: contextual→ Detection: graphs, statistics, testsConsequences: bias of {
coef. estimates
variance est.
predictionscompare,
contrast,
advantages,
disadvantagesRemedies: transformation?or Strategies WLS?

drop variables?

get more data?

add interaction term?

nominal vs. ordinal?

categorical data coding

non-linear models (logistic
(transformation?) regression)?Model assumptionsLecture notes: Chapter 4 section 5Ask for how to detectconsequencesremedies / strategies

Final exam review

Aut 18

2/5

Chapter 5

Indicator variables

nominal vs. ordinal

same effect size for each
"increase" in level ok?

but → more df
flexible (less for
estimating $\hat{\sigma}^2$)
effect for each
category

compare, advantages, disadvantages

Interaction (effect modification)

Lecture notes

Intro 1.2

More details 3

Data Description 1, 1.1

Interaction 1, 3

Chapter 6

Transformations

Lecture notes

Intro: 1.1, 1.2, 1.3,

Summary: 9, 9.1, 9.2

Chapter 7

Weighted least squares

Lecture notes: All 1.2

compare/contrast transformation
vs. WLS

HW 7: Q1

Chapter 8

Correlated errors or autocorrelated ^{over time}

Lecture notes

Diagnostics: 2, residual plot,

4.4 DW: 1st-order autocorr.

4.5 statistic

4.6 interpretation

Transformation: Cochrane-Orcutt

↑ write out w/ 2 predictors
1.1, 1.2

search for "Warning"

Omitted variables

Just 1st sentence

Seasonality

just an example
of an omitted variable

Note: residual plot for ski sales
indicates 2nd-order
autocorrelation,

Note: DW statistic unable
to detect this pattern

Chapter 9

Collinearity

2 predictors

3 or more predictors

or multicollinearity

Lecture notes

Issues: 1, 1.1

Detection: All

Remedies: All

Chapter 11

Variable selection

Intro: All (AIC, BIC)

Modeling Purpose: encourage to read

Automation: briefly describe
forward
backward
stepwise

Chapter 12

Logistic Regression

Just ask: What is response variable?
Why not use linear regression?

Confounding
See Prof. Chen's
notes