

STAT 411/511-2 (Spring 2018) Final Exam Topics

Chapter 1: Statistical conclusions

- 6 steps of Scientific method - where does statistics fit in?
- 6 steps of hypothesis testing
- Scope of Inference
- Randomization and permutation distributions - and why would you use one of these instead of a t -test?
- Sampling distributions

Chapter 2: t -tools

- Central Limit Theorem
- z and t CIs and tests and assumptions for paired, one-sample and two-sample scenarios
- Why would we pool standard deviations from two groups?

Chapter 3: t tool assumptions

- Robustness and resistance of t -tools
- How to check assumptions?
- What are the possible effects of lack of independence on validity of the t -tools?
- How to deal with outliers?
- Box-Cox transform

Chapters 5 and 6: ANOVA

- Why and how do we use information from all samples even though we may only be interested in the difference in means between two groups?
- How to calculate $\hat{\sigma}^2$ from computer output?
- ANOVA hypotheses being tested.
- Be able to compare two models using an Extra SS F-test; hypotheses being tested.
- What are residuals? How to use residuals to check assumptions?
- Robustness and resistance of ANOVA
- PERMANOVA and Kruskal-Wallis - why would you use one of these instead of ANOVA?
- linear combinations
- planned and unplanned comparisons, data snooping
- Individual and familywise confidence and significance levels, Tukey's and Bonferroni tests and CIs

Chapters 7 and 8: SLR

- Interpolation and extrapolation
- How are residuals defined?
- How to calculate $\hat{\sigma}^2$?
- CIs and t tests of regression coefficients, mean response; PI of a future response
- What is *correlation* and when can you say that two variables are correlated?
- Assumptions for SLR and how to check them using appropriate plots
- Assess from a scatterplot whether transformation(s) might be appropriate (Display 8.6)
- Interpretation of results after log transformation of the response variable or the explanatory variable
- Lack-of-Fit F-test

Chapters 9 and 10: MLR

- Interpreting interactions: between predictors; between a predictor and a factor.
- Use of dummy/indicator variables to model categorical explanatory variables.
- Model fitting process (e.g., Display 9.9).
- overfitting
- What are a matrix plot and a Trellis plot and why are they useful?
- Why do coefficients change when explanatory variables are added into the model?
- Robustness and resistance of MLR
- CIs and tests of regression coefficients
- CIs and tests of a linear combination of β_j 's
- What is the Residual SS for a specific model and how do you find it in R?
- How to calculate an estimate of σ^2 ?
- What is the default reduced model used for comparison in the F-test spit out by `lm()` when you fit a MLR?
- Compare two models using an Extra SS F-test; hypotheses being tested.
- How should R^2 be used (and not used) in the context of MLR?

Chapter 11: Case influence statistics

- Fit a rich tentative model then assess Leverage, studentized residuals, Cook's Distance for potential influential points
- How to determine influential data and whether to drop data and how to limit Scope of Inference

Chapter 12: Variable selection

- What is *multicollinearity* and why does it matter? What does VIF tell you?
- AIC, BIC
- all subsets regression - why is it better than step-wise regression?