

Some sample questions for Stat 378 – Final

Concepts:

1. Why would a confidence region for parameter vector $(\beta_1, \beta_2)'$ be of interest?
2. What is the purpose of a standardized regression? What is the difference between the standardized slope and the non-standardized slope?
3. What is multicollinearity and how can you detect it?
4. For each type of residual explain the rationale for the standardization.
5. Explain why a qq-plot for the residuals helps to check for model assumptions.
6. Give a complete list of the model assumptions for MLRM?
7. Explain the rationale for partial residual plots.
8. Sketch a diagram to illustrate the difference between outlier and influential value.
9. Explain Cook's distance as a measure for detecting outliers?
10. What are the DFBETAs? Purpose?
11. What is the Box/Cox transformation? Box/Tidwell?
12. What is the purpose of weighted least squares?
13. How does WLS improve the estimates for the model parameters?
14. What is the effect on the estimators, if a model has more(less) predictors than the population requires?
15. In the model selection process, how can models be evaluated? Name at least 2 measures. Which test should be used to compare models?
16. How does forward selection work?
17. State the components of a GLM?
18. Specify the components of GLM to make it a logistic regression/MLRM.
19. Explain the principle of Maximum Likelihood Estimation?
20. What is the benefit of ridge regression?
21. What is the purpose of lasso regression?
22. What is the role of cross validation?
23. (This is a long one, but perfect for a review) List all steps and their purpose required for a complete/extensive analysis of a MLRM.