Exam 2 list of concepts

Below you will find a skeleton outline of what we have covered so far. Exam 2 will include everything from the 10/24 lecture onward (all content since exam 1). While not explicitly cumulative, the exam content builds on itself. For the exam you will not need to manually calculate critical values, but need to understand the underlying logic behind why they vary with sample size. For the exam you can have either a half sheet of paper (large notecard) front and back **or** a single sided full sheet of paper. I advise bringing a calculator to the exam, but you will not be required to produce any R code (you might be asked to interpret it or tell me what went wrong in the code).

1. Interpretation of both categorical and continuous variables in multiple regression
   1. Understanding how to test for significance without P-value
   2. How to construct confidence intervals, use critical values, etc to make inference about regression slope
   3. F-test and nested f-tests
      1. Interpretation of rejecting/failing to reject null
   4. Parallel slopes vs interaction model
      1. Know how to interpret both, and to know what kind of model it is based on regression output
   5. Indicator variables
      1. Know how to interpret and what it means to be reference/omitted group
   6. Condition violations, understanding how to identify from diagnostics
2. bootsrapping
   1. Logic of why we might want to bootstrap and how it works
      1. Interpretation of bootstrapped regression
3. Model selection and choosing predictors
   1. Tradeoff of adding more variables vs parsimony
      1. Omitted variable bias
   2. Multi-collinearity
      1. Why is it a problem? What happens to our coefficients?
      2. How to test for multicollinearity
      3. What to do if you find multicollinearity?
   3. Mallows C, AIC, and BIC interpretation
      1. Not going to ask you how to calculate, just understand how to interpret
      2. Know how to compare across multiple goodness of fit measures
   4. Process of building a model either forward or backward selection
   5. Cross validation
      1. Logic of it, steps to use
      2. Testing vs training set
      3. Interpretation for correlation between predictions and observations
      4. Shrinkage
4. Logistic regression
   1. Problems of using linear model for binary outcome
   2. Log odds, odds, and probability
      1. How to transform from one to the other
      2. Advantage of transforming to log odds
      3. Interpretation of each measure
   3. Odds ratio and interpretation
   4. G-stat and deviance
   5. Checking for condition violations
      1. Independence, randomness, and linearity
   6. Be able to interpret a logistic regression output
5. Multiple logistic regression
   1. Model assessment for goodness of fit
   2. Predicted probabilities (what we mean by holding all else constant)
   3. Interpreting models in multiple regression, including parallel slopes and interaction model