

Study Guide
for the M.S. Comprehensive Exam
Applied Statistics (22S:164-165)

The following texts are suggested as references:

Abraham & Ledolter, *Introduction to Regression Modeling*, (Duxbury)

Fox, *Applied Regression Analysis, Linear Models and Related Methods* (Sage)

Montgomery, *Design and Analysis of Experiments* (Wiley)

Oehlert, *A First Course in Design and Analysis of Experiments* (Freeman)

Weisberg, *Applied Linear Regression* (Wiley)

Topics:

1. The simple linear regression model
 - a. model assumptions
 - b. least squares estimation
 - c. properties of least squares estimators
 - d. partitioning the total variation
 - e. hypothesis tests on slope and intercept
 - f. confidence intervals for slope, intercept, mean response, predicted value
 - g. regression through the origin
 - h. coefficient of multiple determination (R^2)
 - i. simultaneous inference on slope and intercept
 - j. residual plots
 - k. normal probability plots
2. Multiple linear regression model
 - a. all of the same subtopics listed for the simple linear regression model
 - b. joint, partial, and sequential F-tests
 - c. F-tests for the general linear hypothesis
 - d. polynomial regression
 - e. F-test for lack-of-fit
 - f. Multicollinearity and variance inflation factors
 - g. indicator variables
 - h. F-tests for submodels involving indicators and continuous predictors (e.g., parallel trends)
3. Model selection

- a. selection criteria, including PRESS, C_p , AIC, and BIC
 - b. sequential variable selection procedures
4. Residual analysis and influence diagnostics
- a. residual plots
 - b. standardized, studentized, and externally studentized residuals
 - c. added variable plots
 - d. leverage, hat matrix, and properties thereof
 - e. influence diagnostics such as Cook's distances, DFFITS, etc.
 - f. autocorrelation and Durbin-Watson test
 - g. nonconstant variance and Levene's test
5. Single factor, fixed effects
- a. completely randomized designs and the role of randomization
 - b. ANOVA, F-tests, expected mean square, power
 - c. regression approach
 - d. using SAS
 - e. estimation of effects
 - f. multiple comparisons, contrasts, orthogonal polynomials
 - g. residual plots
 - h. transformation of the response variable
 - i. power and sample size
 - j. single factor, random effects: ANOVA, F-tests, EMS, estimation of variance components
6. Factorial designs
- a. factorial structure
 - b. no interaction
 - c. interaction, more than one observation per cell
 - d. interaction, one observation per cell
 - e. polynomials and division of interaction effects
 - f. unequal cell sizes
7. Block designs
- a. role of blocking
 - b. randomized complete-block designs
 - c. Latin Square Designs, replicating Latin Squares
 - d. blocking of factorials
 - e. incomplete blocks, confounding in 2^k designs

8. Complex designs

- a. random effects
- b. mixed models
- c. repeated measures
- d. crossed and nested factors
- e. split plot designs
- f. rules for expected mean squares
- g. variances of means and contrasts
- h. crossover designs
- i. power and sample size
- j. restricted and unrestricted models
- k. EMS and estimation of variance components
- l. PROC MIXED, ls means and differences

9. Other topics

- a. power transformations
- b. weighted and generalized least squares
- c. variance-stabilizing transformations
- d. Box-Cox and Box-Tidwell procedures to obtain transformations
- e. analysis of covariance
- f. unbalanced data, GLM approach, estimable functions
- g. missing data (empty cells)
- h. Tukey's one degree of freedom for nonadditivity