Fox Module 12 Statistical inference for multiple regression

- Confidence intervals
- Hypothesis testing
- Empirical vs structural relations

Read Section 6.2.1, "The multiple regression model," on pages 105-106.

The five assumptions on page 105 and the five attributes of least squares estimators on page 106 are the same as for simple linear regression.

Know equation 6.2 on page 106, and read carefully the explanatory paragraph afterward. The R^2_j is the "squared multiple correlation from the regression of X_j on all the other X's." Think of a multiple regression with two explanatory variables: R^2_2 is the squared multiple correlation from the regression of X_2 on X_1.

Focus on two critical points in this paragraph:

- The error term σ^2_ε decreases if the explanatory variables are orthogonal.
- The variance-influence factor increases in the explanatory variables are correlated.

Read Section 6.2.2, "Confidence intervals and hypothesis tests," on pages 106-110. Focus on the degrees of freedom for the *t*-distribution (n-k-1) on page 106 and the standard error of B_j at the top of page 107 (which follows directly from the previous section).

The example on page 107 is clear. Expect similar questions on the final exam.

Know both forms of the F-test for the omnibus null hypothesis on page 108: one uses RSS and RegSS and other uses R^2. RSS + ResSS = TSS, so the final exam may give various input data (RSS, RegSS, TSS, R^2) and ask for the F-statistic.

Know the analysis of variance table on page 108. The residual mean square (RMS) is the estimated error variance σ^2_E. You use analysis of variance for qualitative factors and for the student project, so learn the definitions in this module.

The F-test for a subset of slopes is hard to grasp, but it is essential for regression analysis. It is tested on the final exam and is used in the student projects.

- Know the two forms of the F-statistic at the bottom of page 109.
- Distinguish between q (the number of slopes being tested) and k (the number of explanatory variables in the full model.

Know the degrees of freedom (q and n-k-1).

Read Section 6.2.2, “Empirical vs structural relations,” on pages 110-112. This section is intuition, not formulas. Know the relation to the bias of the regression equation in the gray box on page 112. Understand the last line in the box: “Bias in least squares estimation results from the correlation that is induced between the included explanatory variable and the error by incorporating the omitted explanatory variable in the error.”