MS Module 16: Regression estimates (overview 3rd edition)

(The attached PDF file has better formatting.)

(Readings from the third 3rd edition of the Devore, Berk, and Carlton text.)

Reading: §12.2 Estimating model parameters

Know equations 12.2 and 12.3. The proof of these expressions is simple and worth knowing, though you won't be asked to prove equations on the final exam.

Know the definitions of S_{xx}, the sum of squares of the X values, S_{yy}, the sum of squares of the Y values, and S_{xy}, the cross sum of squares of the X and Y values. Know the least squares estimates for β_1, β_0, σ^2, σ, ρ, and R^2. Know the standard deviation of the least squares estimate for β_1.

Know Expressions 12.2, 12.3, and the definitions and formulas right afterward. Example 12.4 shows what you may be asked on the final exam.

Read the section "Residuals and Estimating σ." Know the definitions for the error sum of squares (residual sum of squares) SSE, the least squares estimate of σ^2. If the final exam problem gives an observed *x,y* pair, it may ask for the residual at that point. Example 12.6 shows how to compute all the regression values from observed data.

Read the section "The Coefficient of Determination." Make sure you distinguish between SSE, SSR, and SST; see Figure 12.11. Example 12.7 shows the computation of R^2 from summary statistics.

The textbook shows how to derive the least squares estimators for β_0 and β_1 by the normal equations and then by the computational formulas using the summary statistics.

- Know how to derive S_{xx} and S_{xy} from the summary statistics.
- Know how to derive β_1 from S_{xx} and S_{xy}.
- Know how to derive β_0 from $\overline{x}$, $\overline{y}$, and β_1.

Know how to derive also the estimates of σ^2, σ, R^2, and ρ from the summary statistics. A final exam problem may give summary statistics and ask for β_0, β_1, σ^2, and R^2.

Review the practice problems on the discussion forum. Some final exam problems give X values as the integers from 1 to N, from which you derive Σx_i and Σx_i^2.

The Excel Analysis ToolPak has a Regression module that calculates most of the values taught in ths course. The ToolPak is particularly useful for ANOVA and for regression analysis, since the numerical computations are time-consuming. Other statistical software products are even more comprehensive but are not always available to actuaries.

For efficient study, enter the data from a textbook example into a spread-sheet and run the Analysis ToolPak on the data. Compare Excel's solutions to the textbook solutions to ensure that you are using the ToolPak correctly. Then solve the end of chapter exercises and check your answers with the Analysis ToolPak.

Review end of chapter exercises 13, 14, 15, 16 c and d, 17 b, c, d, and e, 18, 19, 20, 22 b, c, and d.