MS Module 11: Single-Factor ANOVA (overview 3rd edition)

(The attached PDF file has better formatting.)

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

Reading: §11.1 Single-factor ANOVA

The concepts and the procedures introduced in this module are also used in two-factor ANOVA, regression analysis, and actuarial risk classification, covered in later modules. *F* tests are used in many of the modules: focus on the treatment sums of squares in the numerator, the error sum of squares in the denominator, and the degrees of freedom for each.

Treatments are also called groups; dimensions are sets of treatments. The readings on actuarial pricing use the term *classes*.

Know the definitions of the treatment sums of squares, error sum of squares, and total sum of squares and the coefficient to use for each [J vs (J-1)]. Example 11.1 shows the computations.

Know the computational formulas for

- total sum of squares
- treatment sums of squares (for each dimension in two way ANOVA)
- error sum of squares

The final exam problem may give summary statistics to derive sums of squares and mean sums of squares.

Know the format of the ANOVA table (see Table 11.2). The same format is used for two-factor ANOVA and for regression analysis. Final exam problems will not give the means and standard deviations for each row, which you can calculate from the data.

Know how to derive the *F* statistic, and be sure you can determine the *p* value from the tables in the appendix. The practice problems use exact *p* values (from Excel); for final exam problems, you interpolate from tables.

The final exam problems will *not* ask you to compute Bartlett's test or Levene's test. You may skip the text from the second paragraph after Figure 11.2 through the end of the section.

Insurance pricing and risk classification are closely related to single-factor and multiple-factor analysis of variance. Insurers set prices by sex, age, location, and other attributes of the insured. Actuaries determine

- whether a classification dimension is a significant predictor of loss costs (mortality, accidents)
- which groups within a classification dimension are significantly different
- how to estimate expected loss costs within each cell of the classification table

Notation for classification tables uses two or more subscripts, depending on the classification dimensions.

Review end of chapter exercises 3, 4, 5, 6, 8, 9, 10, 11, and 12.