MS Module 3: Confidence intervals, *t* distributions, and variances (overview 2nd edition)

(The attached PDF file has better formatting.)

(Readings from the second 2nd edition of the Devore text.)

Read §8.3 Intervals Based on a Normal Population Distribution

Know Equation 8.13 with Figure 8.7; Equations 8.14 and 8.15 with Example 8.11 and Example 8.12; Equation 8.16 with Example 8.13. These equations are used throughout the textbook.

Skip the sections “Tolerance Intervals” and “Intervals Based on Non-normal Population Distributions.”

Know when to use *z* values vs *t* values.

! If the sample size is large, use *z* values and the central limit theorem.
! If the sample size is small and the population is normally distributed, use *t* values and degrees of freedom.

Confidence intervals and prediction intervals are used in several modules. Prediction intervals are wider, as they include the uncertainty caused by the error term. The relation of the formulas for the confidence interval and the prediction interval is similar for all modules, though the parameters of the formulas change.

Review end of chapter exercises 29, 30, 31, 33, 37 a and b, 38, and 39.

This course covers the arithmetic of statistical testing. If all assumptions are valid, the testing gives plausible results. In practice, the assumptions are often invalid. Half of the results presented in many social science fields are bogus, since the researchers use faulty assumptions. Most commonly, they omit numerous variables that affect the results. For instance, hundreds of papers conclude that police are racist, because blacks are more likely to be arrested than whites. But arrest rates depend on a host of factors, such as felony rates; the conclusions about police racism are usually unwarranted.

The textbook often discusses assumptions, but the final exam does not test these sections. For actuaries, the assumptions are critical to good work. An actuary may find that urban residents have more auto accidents than do rural residents. But urban residents may be younger, wealthier, better educated, less likely to be married, have fewer children; unless one adjusts for these other variables, results may be skewed.

The textbook gives the more exact equation (8.15) instead of equation (8.16), which is the traditional Wald confidence interval, and it explains why the more exact equation is better. Confidence intervals for proportions are used in later modules, based on the reasoning behind equation 8.15. Review example 8.13.

Read the section on “Sample Size Determination.” You need not know equation 8.17, but later modules discuss minimum sample sizes.

Review end of chapter exercises 43 and 44.

Reading: §8.4 Confidence Intervals for the Variance and Standard Deviation of a Normal Distribution

Confidence intervals for the variance of a population use the χ^2 distribution. This distribution is not symmetric, so the confidence interval is not symmetric about the point estimate. Know equation 8.17 with example 8.14.

Review end of chapter exercises 44, 45, and 47b.

Skip §8.5 Bootstrap Confidence Intervals

Bootstrapping is especially useful when the underlying population distribution is not normal, but it requires high computing power and is not tested on the final exam.

Insurance regulation is concerned with extreme value distributions: hurricanes, epidemics, market crashes. The formulas for normal distributions do not apply to these risks, but insurers and financial economists have data bases of past events. They use bootstrapping to estimate tail values at risk and similar risk measures.