MS Module 1 Background and data visualization (overview third 3rd edition)

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

(The attached PDF file has better formatting.)

Readings: §4.3 the normal distribution; §4.6 probability plots, §7.1 point estimation general concepts and criteria, §7.2 methods of point estimation

Chapters 1-7 are also covered on Exam P. The textbook uses these concepts; final exam problems assume you are familiar with them and tests the material in §4.3, §4.6, §7.1, and §7.2. This course assumes you know the material in

§1.2 Graphical Methods in Descriptive Statistics (especially dotplots and histograms)
§1.3 Measures of center (especially mean, median, and quartiles)
§1.4 Measures of variability (especially σ, σ^2, and the computing formulas; boxplots; outliers) [Know the variance of combinations of random variables, especially the difference of two independent random variables.]

§2.4 Conditional probability (especially the multiplication rule and Bayes' theorem) [Note: final exam problems will not ask you to compute results from Bayes' theorem.]
§2.5 Independence

§3.1 Random variables
§3.2 Probability distributions (especially parameters; the cumulative distribution function)
§3.3 Expected values (means and variances)
§3.5 Binomial distributions (distribution function; means and variances)
§3.6 Poisson distributions (distribution function; means and variances)

[Some examples in later chapters of the textbook use binomial or Poisson distributions. The final exam for this course does not test these distributions.]

§5.2 Expected values, covariance and correlation
§5.3 Linear combinations (especially of differences between random variables)
§5.4 Conditional distributions (especially regression to the mean)
§5.6 Transformations of random variables

[The final exam for this course covers hypothesis testing of correlations in linear regression.]

§6.1 Statistics and their distributions
§6.2 The Distribution of Sample Totals, Means, and Proportions (especially the central limit theorem and the law of large numbers)

Most students taking this course already know this material, and the final exam for this course does not test this material. If you are not familiar with this material, or would like a rigorous review of it, review the sections of the textbook listed above.

Read §4.3, "the normal distribution," excluding the last subsection (the normal moment generating function). You are not responsible for the proof that the normal curve satisfies the requirement that its integral is one. This course uses the percentiles of the standard normal distribution, *z* values, *z* critical values, standardized residuals, and normal approximations to other distributions.

Read §4.6, "Probability plots," subsections on "sample percentiles" and "a probability plot," and the two supplementary readings on the discussion forum (see below).

Read §6.3 The χ^2, *t*, and *F* Distributions (general concepts, not the mathematics of the χ^2 distribution, the *t* distribution, and *F* distribution). We use these distributions throughout the course (in addition to the normal distribution).

Read §7.1, "Point estimation - general concepts and criteria," subsections on "mean squared error," "unbiased estimators," "estimators with minimum variance," and "reporting a point estimate: the standard error." The final exam tests the bias, variance, standard error, and mean squared error of estimators. You are not responsible for the subsections on "more complications" and "the bootstrap."

Read §7.2 "Methods of point estimation." Estimation is the core of mathematical statistics, and these methods are tested on the final exam.

Probability plots are used throughout the textbook and they are tested on the final exam. The textbook section of probability plots is brief. Two readings on the discussion forum (and tested on the final exam) are required:

- The section "Normal Quantile Plot" from a reading on *Univariate Analysis*. Know the shapes of the normal probability plots for the four distributions in Figure 3.8. You are not responsible for the Lilliefors confidence bounds in the first graphic.
- Yunsi Wang, Tyler Steele, and EVA Zhang on the "QQ Plot," pages 1-7 only (until the section titled "QQ plot application").

The attachments to this posting have the supplementary readings.

The final exam tests four types of distributions:

- heavy tailed, such as *t* distributions with few degrees of freedom
- light tailed, such the uniform distribution
- right skewed (positively skewed), such as lognormal, Poisson, gamma, and exponential distributions
- left skewed (negatively skewed), such as capped distributions

Look at Figure 8.6, a normal probability plot. A final exam problem may give a probability plot and ask why type of distribution it comes from. These plots occur throughout the text (Figure 8.6, Figure 8.10), generally to show normal distributions. The final exam problems will show various types of distributions.

Know the attributes of the normal distribution, the χ^2 distribution, the *t* distribution, and the *F* distribution. This course does not test the mathematical attributes of these distributions, but you should know their properties to understand how the statistical concepts. You will use these distributions throughout the course; know what each one looks like, and the ranges of their x- and y-axes.