INSTRUCTOR NOTES FOR M358K FOR PART IV (RANDOMNESS AND PROBABILITY: CHAPTERS 14-17) OF DEVEAUX, VELLEMAN AND BOCK, *STATS: DATA AND MODELS*, 3RD EDITION

General comment: Remember that M362K or an equivalent course in probability is a prerequisite for M358K, so this part should be considered mostly review. Do not spend time giving proofs of standard probability properties – that should have been in M362K and is not a goal of this course. At the same time, don't assume students "already know it;" most of them do need some review and may need some adjustment to seeing probability in the context of statistics.

Chapter 14:

- I suggest assigning this chapter as reading (with pp. 342 – 347 "as needed" for individual review), together with some exercises (see below for suggestions), then giving some supplementation in class as indicated below.
- You might want to do a simulation such as the one on p. 339 in class.
- Although this textbook is much better than most at pointing out the differing perspectives/models of probability, I would still give a brief lecture augmenting what is in the text, along the lines of what is in (http://www.ma.utexas.edu/users/mks/M358KInstr/WhatIsProbability.pdf), which can be used as a pdf slideshow for such a lecture. Be sure to point out that this course is about "frequentist" statistical methods – that is, those that use the long-term relative frequency perspective of probability. These methods are not adaptable to the personal probability perspective; Bayesian methods (not included in this course) are better at accommodating personal perobability.
- p. 340 Stress the comment after the box.
- Students should be able to do examples such as those on pp. 342 – 348 from M362K; don't spend time lecturing on basic probability methods, but instead assign some such problems (preferably including some to be graded), with the understanding that they are intended as review of M362K material, and students should consult the textbook for individual review if needed.
- Exercises (pp. 350 – 354): I suggest assigning #1, 9, 13, 15, 17, and 33 for self-check (along with the assignment to read the chapter), spending a little class time taking any questions on the assignment, then assigning #36 (which has several parts requiring different techniques) to hand in the next class day.

Chapter 15:

- pp. 355 – 359 should be familiar material from M362K, so don't lecture on it, but assign it for students to read as needed to do exercises.
- pp. 359 – 365 Students might not come out of M362K with as good an understanding (from the point of view of statistics) of conditional probabilities and independence as they have of basic probability calculations, so these pages warrant a little more attention than pp. 355 - 359.
    - If you have already brought in conditional probability in connection with Chapter 3, you may need to spend less time on it now.

  - Some students may benefit from the disk activity Birthweights and Smoking.
  - Don't spend time doing proofs here – these should have been seen in M362K. In M 358K, the emphasis needs to be more on recognizing the concepts in context.
- pp. pp. 366 – 367 Students should have seen drawing without replacement in M362K, but the connection with finite populations is important here. There will be more on this subject in Chapter 18.
- pp. 367 – 369 Many students will have seen tree diagrams in M362K, but some may have not, so this section is worth spending a little time on – although perhaps just to the point of assigning the section as reading and some homework.
- pp. 369 – 373
  - Students should have discussed Reversing the Condition in M362K. The important thing to emphasize here is not to confuse P(A|B) and P(B|A).
  - I would omit the example on p. 369 (and not assign similar exercises), since students should have done such examples in M362K; although such examples are important, they are not really the subject of M 362K.
  - It might be worthwhile to go through the Math Box on pp. 372 – 373, for the benefit of student students who have not used tree diagrams before.
- p. 374 The "What Can Go Wrong" points are worth emphasizing.
- Exercises (pp. 275 – 279) Suggestions: #1, 3, 5, 7, 9, 15 as self-check; #12, 26, and 38 to hand in.

Chapter 16:

- This chapter is mostly review of material from M362K, so assigning most as reading with exercises for self-check and/or class discussion, perhaps supplementing with short lectures on some places, is probably the best way to proceed.
- Students may have seen a more formal, mathematical definition of random variable in M362K, but the one given on pp. 381-382 ("numeric values based on the outcome of a random event") is more helpful for applied statistics. (For a variant of this that I usually use, with some examples that point out common sources of confusion, see Random Variables, http://www.ma.utexas.edu/users/mks/M358KInstr/RandomVariables.pdf, which may be used as a handout, short lecture, or just for examples or ideas.)
- On p. 282, the authors state, "We use a capital letter, like X, to denote a random variable. We'll denote a particular value that it can have by the corresponding lowercase letter, in this case x." I try to follow this practice, since it helps emphasize the difference between a random variable and a value of the random variable, which is helpful in dealing with the sampling distribution. Unfortunately, in Chapter 18, the authors do not consistently follow the practice they state here. On p. 434, they mix practices, using capital Y for number of successes but small p-hat for proportion of successes. But by p 442, they seem to have abandoned the practice of using capital letters for random variables entirely: The use small y-bar for the random variable sample mean, and on p. 443, in the Math Box, they talk about Var(y_1), etc.

- p. 384 Remind them of the M362K definition Var(X) = E((X – μ)^2), and show how that coincides with the use of variance here, where the random variable is discrete.
- pp. 386 – 389 The formulas here should be familiar from M362K.
    - Be sure to emphasize how the formulas for variance are different from those for the mean.
    - Emphasize that the formulas for the variance of sums and differences of random variables apply only to independent random variables.
    - The box on p. 389 is important.
- pp. 390-394 Emphasize the items on p. 390 (The Bad News) and 392 (The Good News), since these may have slipped under the radar in M362K, but are important for understanding sampling distributions.
- pp. 395-396 Have them read the starred section; it should be mostly review from M362K. The most important thing to emphasize:
  Var(X +Y) = Var(X) +Var(Y) + 2Cov(X,Y).
- As usual, the "What Can Go Wrong" section (pp. 396 – 397) is important.
- Suggested Exercises (pp. 399 – 403): 19, 27 for self-check; 38 and 40 to hand in

Chapter 17:

- This chapter is somewhat awkward for this course. Bernoulli trials and the binomial distribution are important for introductory statistics, but this chapter introduces them via an example using the geometric distribution, which is not relevant to the course material. One approach would be to assign the chapter as reading, but tell the students that the important parts for this course are Bernoulli trials and the binomial model; they don't need to be concerned with the geometric model for this course, except to be able to distinguish between the binomial and geometric models. Students should have seen Bernoulli trials, the binomial distribution (including the formulas for its mean and variance, and proofs as in the math box on p.410), the geometric distribution, and the normal approximation to the binomial (p. 412) in M362K, so pp. 404 – 412 should be considered review.
- p. 407 The 10% Condition may be new. It will be discussed more in connection with Chapter 18. You may want to look at the supplement "Where Does the 10% Condition Come From?" now and decide whether to use it now, or in Chapter 18, or split it up.
- p. 412 There are a number of web demos showing the normal approximation to the binomial that you may prefer to the one on the ActivStats disk. Several are linked from the External Links section of http://www.ma.utexas.edu/users/mks/M358KInstr/M358KInstructorMaterials.html
- The Success/Failure Condition (p. 413) is likely to be new, and is important for applications and as an example of other "appropriateness of approximation" conditions, so it is worth going through in class (including the derivation in the Math Box, and the footnote).
- pp. 414 – 417 Omit the Poisson and Exponential models. Students should have studied these in M362K, and they are not used in M358K.
- p. 419 The "On The Computer" section may be worth assigning as reading.

- Exercises (pp. 421 – 424) Possible exercises: 1, 7, 17 as self-check; 24 and 30 to hand in.

Review Exercises (pp. 426 – 430) Possible assignment: #2, 3, 10, 24

Last Revised Oct. 14, 2012