Vanderbilt University
Political Science Department
Fall 2023

Stats 1
(PSCI 8356)
Professor Jim Bisbee

PROBLEM SET 1: Due Tuesday, September 12th at beginning of class.

A reminder: you may work with others in the class on this problem set, and you are in fact encouraged to do so. However, the work you hand in must be your own. Handwritten work is acceptable, but word-processed work (e.g., using LaTeX or RMarkdown) is preferred.

1. On the class GitHub site may be found a link to a dataset containing information on college undergraduate programs in the United States. In a *very* brief, two-paragraph mini-essay, describe the data therein. In your discussion, you must correctly use each of the following 8 terms. For Martin's sake, please **underline** the terms as you use them in your mini-essay. Don't worry too much about the content or flow of your essay: I just want to see you using these terms correctly. **NB:** You will need to *look* at some of these variables by visualizing them in R.

interval-level variable	a distribution skewed to the right
nominal-level variable	mode
units	dichotomous variable
symmetric distribution	median

   The following two questions are designed to get you nimble with the sorts of proofs we'll be doing in class with scalar algebra and summation signs:

2. Prove the Law of Total Probability, that is:

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

3. Show that the probability of a set A of equiprobable simple events E_i is equal to the number of simple events in A over the total number of simple events in S, that is:

$$P(A) = \frac{|A|}{n}$$

   *Most of the remaining problems come from WMS, **7th edition**, which is on our Github. Note that the answers to odd-numbered, non-asterisked problems may be found in the back of the book. (In contrast to those answers, please show all your work.)*

4. WMS Exercise 2.124

5. WMS Exercise 2.125

6. WMS Exercise 2.47

7. There are 16 students in Stats I. Of these students, 11 are male. I assign the students at random to three teams consisting of 6, 5, and 5 students respectively. Figure out the probability of having all single-sex teams.