

## Writing activities

These notes come from the Chance Instructor's guide.

For more information on the Chance project, see:

<http://www.dartmouth.edu/~chance/>

We have seen that one of the formats in which Chance has been successfully offered is the Freshman Writing Seminar. The curricular goal of such courses is to provide first-year undergraduates with a small, discussion-based (as opposed to lecture-based) course where there is significant focus on the writing process. We have found the Chance course to provide an ideal context for such a seminar.

Writing in statistics courses is in general gaining respectability. Gudman Iversen ("Writing Papers in a Statistics Course." 1991 *Proceedings of the Section on Statistical Education* -- American Statistical Association (1991):29-32) in rationalizing his use of writing in his statistics courses says, "... any writing is good for you." Norean Radke-Sharpe ("Writing as a Component of Statistics Education. *The American Statistician* 45(1991):292-293.) gives (and then expands upon) advantages of required writing in a statistics course:

- (1) it improves writing skills,
- (2) it focuses internalization and conceptualization of material,
- (3) it encourages creativity, and
- (4) it enhances the ability to communicate methods and conclusions.

In her paper Radke-Sharpe gives a variety of suggested writing assignments.

Many of the technical terms from probability and statistics have counterparts in everyday speech, where they are used less carefully than would be the case in scientific discourse. It is obvious that people frequently talk about chances, odds and likelihood without relying on a formal probability model. But also, people invoke a so-called "Law of Averages", which is often used to defend conclusions that don't follow from the statistician's Law of Large Numbers (see the article on the "Law of Averages" by Ann Watkins in the Spring 1995 edition of *Chance* magazine). Similarly, most people have a nodding acquaintance with the idea of a "bell curve", without understanding conditions under which appeals to the Central Limit Theorem might make sense, or when other data models might be appropriate. Writing assignments requiring the exposition and application of such concepts offer valuable lessons on the precise use of language. Students find they must find a balance between using their own voice and the risk of blurring sometimes subtle technical distinctions.

Journals may be useful in the seminar as informal or "free writing" opportunities to record thoughts on the day's discussion in class, pose questions for the instructor, and record solutions to homework exercises. However, the seminar format also requires more structured writing assignments.

We have found that the topics in a Chance course lend themselves to a number of different writing formats. Listed in the appendix are details and comments on writing assignments from two different Chance courses.

### **Comments by Bill Peterson on writing assignments.**

Recall that the writing assignments for my course at Middlebury were:

<i>Style</i>	<i>Assignment</i>
Narrative	Describe a personal experience where you think statistical knowledge would have been helpful to you.
Expository	Describe for a non-specialist the key ideas involved in obtaining information about a population by sampling, citing examples from current news stories.
Critique	Debunk three examples of graphical abuse that you find in the popular press.
Analysis	Simpson's paradox (pitfalls of aggregation) in data on race and imposition of the death penalty.
Argument	Should there be mandatory screening of health care workers for HIV infection?
Research (final project)	Research paper on your choice of Chance topic in the news.

The first five assignments here ranged between 1 and 4 pages in length; the last was an 8-10 page research paper. The narrative paper was assigned on the first day of class, to get the students immediately involved in writing.

The expository paper on polls proved surprisingly difficult for the students. Formulating a clear statement of what a "margin of sampling error" means is not an easy task: it brings to the fore exactly the issues described above regarding colloquial connotations vs.

technical meanings of key terms. Critiques of newspaper reports were illuminating here. The *New York Times* commonly includes a side-bar with an article describing a sample survey, entitled "How the Poll was Conducted." There one finds statements such as "In theory, in 19 cases out of 20 [the 95% confidence statement] the results based on such samples will differ by no more than five percentage points [the margin of sampling error] in either direction from what would have been obtained by seeking out all voters in the country." In one instance, it was reported that this "latest poll was conducted by telephone interviews...The sample of telephone exchanges was selected by a computer from a complete list of exchanges in the country. The exchanges were chosen so as to assure that each region of the country was represented in proportion to its population [stratified sampling]." Other polls are less clear in reporting their methods or the proper interpretation of their results. A perhaps disturbing number are now including disclaimers of the form "This is not a scientific poll, but *represents a sampling* of reader opinion...[emphasis added]". It would appear that, even as they confess their use non-probability sampling, these polls drop terms that attempt to sound like the statistician's goal of a "representative sample". Sorting out language issues like this is an important exercise.

The critical paper on graphs was adopted in the spirit of Edward Tufte's classic *The Visual Display of Quantitative Information* (Cheshire, Connecticut: Graphic Press, 1983). The idea is to contrast the exploratory data analysis techniques introduced in the course to the grossly misleading data displays so often found in the popular press. As with the polling paper, this led to interesting class discussions on why these abuses persist.

The analysis paper on Simpson's paradox was an eye-opener for many students. Inevitably, the first presentation of an aggregation paradox in class leads to a stunned silence. The phenomenon sheds light on what students recognized as a common occurrence in political debates: both sides are able to quote summaries of the same data that appear to support their own point of view. It is instructive to see that there is a rational way out of the paradox, and that one is not forced to abandon hope and conclude that "you can prove anything with statistics."

The assignments grew more involved as the course progressed. In the paper in HIV testing, students were required to take a stand on a controversial issue (Kim Bergalis' compelling testimony before Congress was in the news at this time), and to assemble evidence to support their position. The final paper required them to bring all of the analysis ideas from the course to bear on a story that they found personally interesting. In addition to the statistics techniques, we had scheduled classes with the library on research techniques, including electronic database searching.

### **Tom Moore's Writing Assignments**

Here is a second set of assignments, for a different offering of the course (Moore, Grinnell, Fall 1993). This perhaps comes closest to the original conception of the

course, which was to select five major topics for the syllabus. Here each topic gives rise to a writing assignment.

<i>Course Unit</i>	<i>Writing Assignment</i>
I. Public Opinion Polls	Find and critique the reporting of a poll by the press using the outline on p. 41 of Moore's <i>Concepts and Controversies</i> book.
II. Clinical Trials and Other Kinds of Studies	Find an article from the popular press that is about some new scientific study. Briefly describe the study in the new article. Discuss the extent to which the article tells you enough to assess the validity of the study, and generate a list of questions that would help you assess validity but are <i>not</i> answered by the article.
III. Coincidences	Describe a coincidence in your own life and why you think it is a coincidence and what the likelihood of such an occurrence would be.
IV. Data Analysis and Reporting Numerical Information	Find and critique a graph in the news.
V. Deming and Quality	Based on your brief tour of Company X (the class had gone on a field trip to a local area business), what would be Deming's three primary suggestions for improvement?

### **Additional Resources on Writing**

Connolly, P. and T. Vilardi. *Writing to Learn Mathematics and Science*. New York: Teachers College Press, 1989.

Iverson, Gudmund R. "Writing Papers in a Statistics Course." 1991 *Proceedings of the Section on Statistical Education* -- American Statistical Association (1991):29-32

Radke-Sharpe, Norean. "Writing as a Component of Statistics Education. *The American Statistician* 45(1991):292-293.

Sterrett, Andrew (ed.). *Using Writing to Teach Mathematics*. MAA Notes, Number 16.

Zinsler, William. *On Writing Well: An Informal Guide to Writing Nonfiction*, 3<sup>rd</sup> Edition. New York: Harper & Row, 1985.

