

TITLE: Coupling Literature and Statistics

AUTHORS: Robert Boruch and Judith Zawojewski

REPORT NO.: A-415

PUBLISHED

IN: Teaching Statistics, 1988, 10(2), 37-42
and
D. Green (Ed.) (1994) Teaching Statistics at its Best. In press.

Coupling Literature and Statistics

Robert Boruch and Judith Zawojewski
Illinois, USA

INTRODUCTION

Did you know that Alexander Dumas wrote a short story entitled *Statistics*? Or that *Robinson Crusoe*, *Gravity's Rainbow*, and innumerable other stories often exploit simple ideas about probability and statistics? Are you aware that humorists, such as Mark Twain, fiction writers such as Len Deighton, novelists such as Thomas Pynchon, and autobiographers such as Ben Franklin, present ideas (rightly and otherwise) based on the gamblers' fallacy and simple distributions?

Most people are not aware that a great deal of probability and statistics can be found in prose, poetry and drama. Smeeton and Smeeton (1984, 1985) illustrated ways in which statistical ideas can be introduced in many facets of English Studies, including literature. Their studies have indicated the feasibility of blending statistics and English Studies. It seems timely, then, to share a collection of illustrations in literature and history that have been compiled by one of the authors over a number of years. They are based on readings that are common in precollege and liberal arts college curricula.

ILLUSTRATIONS IN HISTORY

The 17th and 18th centuries constitute a watershed period for thinking about probability and statistics. The works by Graunt and others on civic and medical statistics are rich examples. Poisson's work as a military statistician for Napoleon was disseminated and adopted in 18th century philosophical writing. That in turn was adopted during the precolonial and colonial periods in the United States. To judge from Garry Wills' *Inventing America*, Jefferson's constitutional writing and personal papers contain strong influences of numerical thinking characteristic of the earlier period. Madison's *Federalist* papers also exhibit the themes.

The 19th century growth of applied statistics appears clearly in popular literature of the period. There were similarities in some respects to contemporary debates. Assaults on the validity of census data, for instance, worried journalists, businessmen, and politicians, as well as census officials. Counts were all important in government representation, government subsidy of geographic exploration, business, and so on. Even the subtler matters appear early. The conflict between two cultures about which C.P. Snow has written was

explored in the 1890's by both Italian and American statisticians.

ILLUSTRATIONS IN LITERATURE

Statistical reasoning, correct and otherwise, is not difficult to find in pre-20th century fiction. For instance, the gambler's fallacy is found in Edgar Allan Poe's *The Gold Bug*, in fiction by Dorothy Sayers, in Arthur Conan-Doyle's *Sherlock Holmes* stories, and in Kahn's *The Code Breakers*. From William Kruskal's thoughtful treatment of numbers in prose, we find Charles Dickens loathing dreadful statistics and Anthony Trollope loving them. Rubin's remarkable series of papers for the *American Statistician* leads us through the statistical worlds of Herodotus, Thucydides, Darwin and William Shakespeare. Some of the same spirit is reflected in more contemporary treatments of scientific evidence by William Golding, author of *Lord of the Flies*, in reviewing the works of Herodotus and H. G. Wells.

Other statistical ideas can be identified in contemporary work. Michael Crichton explores discriminant analysis, for instance, in his recent *Congo*. He alludes to it in *Andromeda Strain* to help determine if his apes are really apes from language, appearance and behaviour. Francis Crick's problem in extracting X-ray photographs from Rosalind Franklin in *Double Helix* has precedent in early scientific ideas about Pythagorean privacy of research (to judge by de Santilla's biography of Galileo). The problem appears in more modern form in popular treatments of scientific work, such as Koestler's *Midwife Toad* and Jeremy Bernstein's *Scientific Experience*. John Cheever exploits primitive notions of coincidence and likelihood in the opening to *Bullet Park*. In *Bomber*, Len Deighton errs in computing simple probabilities for being shot down in missions during World War II.

In more popular writing, science fiction and spy novels are a rich source of examples of the use of statistical ideas. Isaac Asimov, C. Simak, and other science fiction authors refer repeatedly and explicitly to odds, numerical probability, distributions and other concepts. In a recent science fiction television series and book, *The Hitchhiker's Guide to the Galaxy*, improbability is a keystone of the stories.

Teaching Statistics through English Classes

Clearly, there are many illustrations of probability and statistics in prose, poetry and drama. The examples given here are the result of casual interest, or hobby, of collecting such examples. If we are to seriously consider embedding statistics lessons into non-mathematics courses, we need to systematically assess the actual readings, assignments and discussions that occur as a part of regular coursework. We have recently begun such an attempt as part of an National Science Foundation funded project on *Statistical Literacy: A Cross-Discipline Development Project*. At present, we are compiling lists of "typical" literature readings required of ninth grade students in the Chicago metropolitan area. Then, we are systematically reading those sources with the purpose of identifying any uses of probability, statistics or statistical reasoning. Based

on the illustrations already shared in this article, we have reason to believe that we will find many places in which short introductory lessons (or vignettes) in statistics may be embedded. Coupled with the information from Smeeton and Smeeton (1985) that such lessons are feasible, we hope we can begin to integrate instruction in statistical thinking into all areas.

References

- Smeeton, N.C. and Smeeton, E.A. (1984). Statistical Ideas in English Studies. *Teaching Statistics*, 6(3), 92-94.
- Smeeton, N.C. and Smeeton, E.A. (1985). The Interface Between English and Statistics. *Teaching Statistics*, 7(2), 34-37.