STATISTICAL LITERACY - STATISTICS LONG AFTER SCHOOL

Jerry L. Moreno, John Carroll University, USA

Decisions made by citizens, or made by others that affect them in their personal, social, and working lives, are often based on statistical surveys, polls, and experiments. It is therefore important that citizens have some understanding of just how appropriate were the statistics on which decisions are made. Moreover, citizens should know enough statistics to be able to read a newspaper intelligently. A list of statistical concepts for Citizens Statistics 101 is given which could serve as the knowledge base that effects such understanding. The ultimate goal is to create a quantitatively literate citizenry.

RATIONALE TO CREATE A QUANTITATIVELY LITERATE CITIZENRY.

Why should a citizen care to be quantitatively literate? One reason is to be able to read a newspaper intelligently (Gani, 1982). H. G. Wells wrote that "Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write," and David Moore (1990) echoed "Changing trends in most societies will make numeracy as essential as literacy now is." Has that day yet arrived? To many, it most certainly has. The real-world personal influences of advertising, political claims and public policy, business and economic forecasting are often based on data analysis. But to what extent does the average citizen possess the tools to understand these influences and to be able to think, assess, and make decisions statistically? As a simple example, how would he or she interpret the use of the word *average* in the last sentence? Would it be his/her same response to *average* in "the average remaining lifetime of people your age is 27 years?" Would it be "*most* people my age live 27 years more" or "half live more than 27 years more and half live less than 27 years more" or "add up all the years remaining and divide by the number?" When asked what the latter means, would the response be "well, you know, *average*!" Almost certainly the response would not be "if everyone my age were to have the same number of years left to live, it would be 27" or "27 years is the balance between the total number of years for everyone whose remaining years were more than 27 to the total for everyone whose remaining years were less than 27."

The National Council of Teachers of Mathematics (NCTM, 1989) recommended that statistics become an integral part of the mathematics curriculum in grades K-12. With schools adopting the NCTM Standards' recognition that statistics teaches students to question, conjecture, and search for relationships when formulating and solving real-world problems, seeds of statistical literacy are being planted in the elementary grades.

As students advance through the grades, they should be linking their knowledge of statistics to other content areas in the social and natural sciences. By doing so, they will develop an understanding of the scientific method and the concepts and processes used in analyzing data, a necessary knowledge base in order for them to become able to make critical and informed decisions throughout their lives. Think of it! They may even *want* to converse about *quantitative* matters as naturally as they do literary, social, or political topics in casual conversation. Well, if that sounds a bit too idealistic, akin to “bears will use indoor plumbing first!” (from George Cobb), then at least consider it a goal. It would certainly go far in easing Paulos’ (1988) concern that citizens do not even recognize the consequences of mathematical, probabilistic, and statistical ignorance.

CITIZENS STATISTICS 101 - STATISTICS CONCEPTS EVERY CITIZEN SHOULD UNDERSTAND

At previous ICOTS meetings, Kruskal (1986) and Moore (1990) have suggested topics for all citizens. Also, Scheaffer, Watkins, and Landwehr (1997) wrote that every high-school student should know: number sense (via graphs and tables), planning a study/producing data (via experiments and surveys), data analysis (graphically looking for patterns and summarizing key features of univariate data, and looking for association in different types of bivariate data), probability (the study of random events and distributions, definitely not combinatorics), and statistical/inferential reasoning (through confidence intervals). It is an excellent outline for a K-12 curriculum, but it will be a long time before all students attain that level of statistical understanding. Until then, enroll all citizens in *Citizens Statistics 101*.

FUNDAMENTAL IDEAS FOR CITIZENS STATISTICS 101

The reasons for statistics

There are many reasons for the subject of statistics but basic ones would certainly include the omnipresence of variation, making decisions, and making sense of information.

The understanding of graphs

Statistical graphs appear in almost every daily and weekly publication. Huff (1954), Wainer (1997), and many others have written on features of graphs that can mislead the viewer including gee-whiz graphs with missing or arbitrary vertical scaling, pictographs whose picture dimensions do not reflect the respective data proportions, and graphs cluttered with "chartjunk." Citizens should be aware of these pitfalls, understand the features of good graphics, and be able to interpret them.

The purpose of basic summary statistics

Citizens should have a basic understanding of a summary measure whose purpose is to characterize some attribute of a set of data (and which is used later in the course to help make decisions). To make sense of this, an analogy is to think of a person, and characterize that person by *one word*. One word cannot possibly tell us everything there is to know about that person, but a collection of words begins to paint a picture. For example, some words such as tall, strong, athletic, heavy tell us something of the person's physique attribute. Each of these words has a different sense in which the person's physique is described. Other attributes could be intellect, sexual appeal, or personality.

So also is a set of data characterized, not by words, but by numbers. The attribute of *center* of a data set is measured in various senses: most often (mode), physical middle (median), halfway split between lowest and highest (midrange), and balance of total sum of positive deviations with total sum of negative deviations (mean). Other measures are the *spread* of the data set, and perhaps its *shape* (a topic of Citizens Stats 102).

The design of an experiment, observational study, survey, poll

Citizens should know that experiments, surveys, and polls must be carefully planned, e.g., use of randomization, in order to provide scientific information worthy of consideration and to avoid bias of various types that can invalidate conclusions. To understand the features of a well-constructed design will prompt them to ask probing questions concerning recommendations affecting them, whether from advertisers, public officials, or their own physician.

Although polls are read daily, how many citizens understand the underlying principles on which they are based? As an analogy, consider being seated in a large dark room facing a mural covering the opposite wall. You are to determine various attributes regarding the mural, such as its predominant color or topic. The difficulty is that the room lights are burned out, but there are a limited number of spotlights available (very expensive to operate and you must pay for their use). How many spots should you ask for? What locations should you choose? Based on a shockingly small number of randomly chosen spots, you are amazed that you are able to conjecture a general color for the whole mural and have a suggestion or two for its theme. And so Gallup chooses 1500 Americans to accurately conjecture what millions are thinking. Being able to convince the citizen that the random procedure does indeed work is not an easy task. But it is an essential one since the natural reaction to something that is not understood is to distrust it, or to believe only that which supports one's own point of view. How wonderful it would be to witness a citizenry so educated that it would revolt against the charlatanry of self-selection television and web call-in polls (call yyy-yyyy to register a yes vote, nnn-nnnn for no)!

Statistics does not prove anything

A statistical analysis provides probabilistic evidence in favor of one interpretation over another. But to the uninformed mind, this lack of being able to establish a theory with mathematical certainty (proof) is thereby interpreted as being able to establish anything desired by using statistics. *Lies, damned lies and statistics*, so said Disraeli. A criminal courtroom case in which there were no eyewitnesses is an excellent analog to what statistics can and cannot do. The charged person is assumed innocent. Bits of evidence are collected and a verdict of innocent or guilty is concluded based on the evidence. It is not really known whether an innocent person was condemned or a guilty one set free. It is crucial that our informed citizen realizes that whatever the conclusion, the person has neither been *proven* innocent nor *proven* guilty.

Uncertainty and concept of probability: A drink a day is good; a drink a day is bad.

When reading an article, citizens should be aware of whether it is just a progress report on an initial pilot study, or something more substantial. Scientific findings are often probabilistic in nature, and thus, depending on how sophisticated a study is, it may produce a conclusion that is contradictory or not supportive of another. Cohn (1989) has written a wonderful book in which he informs his fellow reporters of the importance of understanding this, and statistics, so that they themselves are better equipped to write about output from scientists, doctors, environmentalists, politicians, economists.

Misleading conclusions implied from newspaper articles headlines

Before believing a headline ("Precocious violinists gain brain advantage" or "AIDS virus is spreading among Ohio women"), the citizen should read the article keeping some major statistical items in mind including: sample size(s) used, how the subjects were selected, general design (a controlled experiment or an observational study), whether or not a relationship can be considered causal, the presence or absence of a control group, extrapolation of conclusion beyond appropriate bounds, misdirection of conclusions, who conducted the analysis and who funded it, the phrasing of a survey question.

"How big is big?" - THE question of statistical hypothesis testing

The statistically aware citizen should have a conceptual understanding of the basic structure of hypothesis testing. Suppose that a new rather expensive nonprescription drug appears on the market that is supposed to increase one's retention of information. A simple study was done based on the number of recalled facts over an extended period of time. What should the citizen understand about hypothesis testing? That subjects are randomly assigned to a control group who are given a placebo, or an experimental group who are given the actual drug. That the same experimental procedure is performed on all, and the number of recalled facts recorded. That a summary statistic (mean or median in this case) is measured on each group. That there is a statistical procedure to determine whether the difference between the summary measurements is "big" enough not to be attributed to chance, so to infer the difference to the population at large as well.

THE MATERIALS OF CITIZENS STATISTICS 101

There are many books and resources that can be used for this course including the local newspaper and Laurie Snell's *Chance* (http://www.geom.umn.edu/locate/chance). A suitable text would be Moore's *Concepts and Controversies* (1997) with support readings from Cohn (1989), Huff (1954), and Paulos (1988).

HOW CAN CITIZENS STATISTICS 101 BE IMPLEMENTED?

In the United States, the American Statistical Association (ASA) should be the organizer of such a project possibly finding funding through a National Science Foundation initiative. Local organization would be through the ASA chapters and community education departments that are located in most school districts, with support from parent/teacher organizations. An activity that might encourage citizens to learn statistics would be to add open categories to the American Statistics Poster and Project Competitions, as the Japanese have. Developing a television series would be terrific!

CONCLUSION

Citizens have a “supermarket tabloid” understanding of statistics. They should be equipped with sufficient statistical tools, even if just conceptual and intuitive, to be able to have a general understanding of decisions and conclusions that are based on statistical evidence. It is hopeful that future generations will be thus equipped through school statistics. The statistics community must create Citizens Statistics 101 if there ever is to be a chance of having a truly quantitatively literate citizenry.

REFERENCES

Cohn, V. (1989). *News and Numbers: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields.* Aimes Iowa: Iowa State Univ. Press.

Gani, J. (1982). The Relevance of Statistical Training. Proc. ICOTS I. Sheffield, 50-68.

Huff, D. (1954). *How to Lie With Statistics* (reissue 1993). New York: W. W. Norton.

Kruskal, W. (1986). Is Anyone Listening Out There? Statistics and the Public. Proceedings ICOTS II. University of Victoria, 33-37.

Moore, D. S. (1990). Statistics for All: Why, What, and How? Proceedings ICOTS III. University of Sheffield, 423-428.

Moore, D. S. (1997, 4th ed.) *Statistics: Concepts and Controversies*. NY: Freeman.

National Council of Teachers of Mathematics (1989) *Curriculum and Evaluation Standards for School Mathematics.* NCTM, Reston, Virginia.

Paulos, J. A. (1988). *Innumeracy: Mathematical Illiteracy and Its Consequences*. New York: Hill and Wang.

Scheaffer, R., Watkins, A. Landwehr, J., and Lajoie, S. (Ed.) (1997). Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12. Chapter 1. Hillsdale NJ: Lawrence Erlbaum Associates.

Wainer, H. (1997). *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. New York: Copernicus, Springer-Verlag.