# TOWARD A STATISTICALLY LITERATE CITIZENRY: WHAT STATISTICS EVERYONE SHOULD KNOW

Jerry L. Moreno
John Carroll University
USA

*In my paper 'Statistical Literacy - Statistics Long After School' presented at ICOTS V in Singapore, I discussed creating a course 'Citizens Statistics 101' and suggested what topics should be included in such a course. Unfortunately, the project has not happened, but I continue to think about it and present in this paper a practice exam that illustrates its content. The bottom line hope is that someday there will be no need for such a course as students will learn the statistics in school that will enable them to be statistically literate citizens.*

## INTRODUCTION

What is meant by *statistical literacy*? In the 1930's, John Dewey (cited in Steen, 2001) spoke of literacy as *popular enlightenment* meaning that which enables people to think for themselves, judge independently, and discriminate between good and bad information, as opposed to the passive literacy of being able to understand instructions and carry out procedures routinely. Applying this concept to statistics makes sense to me. Surely, a citizen should be able to read a newspaper intelligently, make decisions based on logic and quantitative information regarding political candidates, medicines and health, investments. Practically everyone in the workplace from farmers to lawyers, jurors to the accused, manufacturers to consumers needs to be able to think quantitatively. But Paulos (1988, 1995) indicates very well that citizens are not so equipped, and Steen (2001) warns that "an innumerate citizen today is as vulnerable as the illiterate peasant of Gutenberg's time." We are trying to solve the problem for future citizens by improving the teaching of statistics in our schools so that for them quantitative reasoning is as commonplace and demands the same respect as do reading and writing. But what can we do to enrich the lives of our current citizens?

## FUNDAMENTAL IDEAS OF CITIZENS STATISTICS 101

I suggested in my ICOTS V paper (Moreno, 1998) fundamental statistical ideas that citizens should appreciate. They were: the reason for statistics; the understanding of graphs; the purpose of basic summary statistics; the design of an experiment, observational study, survey, poll; statistics does not prove anything; uncertainty and the concept of probability (including risk analysis); misleading conclusions implied from newspaper articles headlines; and "How big is big?" - *the* question of statistical hypothesis testing. It was suggested after the talk that I should have also included indexes. I agree that indexes is important for a statistically literate citizen to understand although it is a topic that I dare say few of us teach in an introductory statistics course, except possibly in business stats. What exactly does the Consumer Price Index measure? It is used to measure inflation, but does it, and if it does, how does it? Other indexes that impact society are Producer Price Index, Consumer Confidence Index, Human Development Index, Gas Price Index, Health Care Quality Index, etc.

## A POSSIBLE EXAM FOR CITIZENS STATISTICS 101

Assuming general agreement with the fundamental ideas listed above, what might a take-home examination look like for this course? These questions are based on real situations found on the Internet or in books or the newspaper.

1. According to the United States Department of Agriculture, the average American consumes 20 teaspoons of sugar per day. In the context of this statement, explain four possible interpretations of the word "average" in the USDA's statistic. Which do you think they meant and why?

2. John Lott, senior research scholar in the School of Law at Yale, has conducted an incredibly extensive study of crime deterrence and has concluded *More Guns, Less Crime*, the title of his book. Is he right? Summarize his point of view. Was his study an observational one or an experiment? (See Lott, 2000).

3. Suppose that you are on a jury that is to decide whether the Ford Explorer is a dangerous vehicle to drive or its Firestone radial tires have inherent defects. Note that testimony to date indicates the test rig used by Ford showed that the Firestone tires under suspicion tear apart whereas Goodyear tires do not. The Firestone study subject to the Society of Automotive Engineers specifications indicates that Firestone is the same or better than other tires. Who's right? What questions would you ask?

4. Suppose that you are in charge of determining whether or not a space shuttle launch should occur. The temperature for tomorrow's launch is predicted to be 31°F. You are aware that there is a question regarding the possible effect of low temperature on O-ring performance. O-rings seal the joints between different sections of the solid rocket motors. a) What questions would you ask before making a decision to launch? b) Data are available from 23 previous space shuttle launches giving the temperature at launch time and the number of incidents of O-ring damage in that launch. Seven of the flights had experienced some O-ring damage. The scatterplot of these seven flights is shown in Figure 1. What decision would you make based on this scatterplot? c) Is there anything else you would like to know before you make a decision? d) Make a decision and provide reasons. (see Landwehr & Watkins, 1995, pp.144-146.).

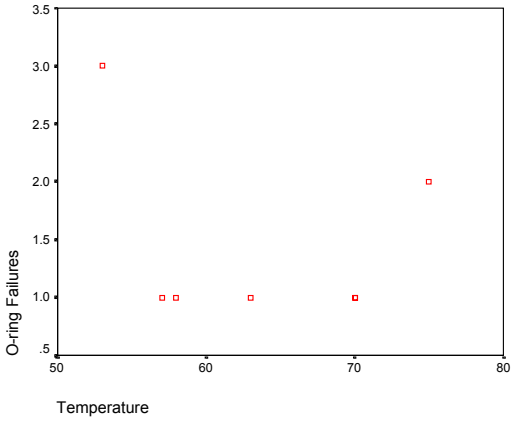
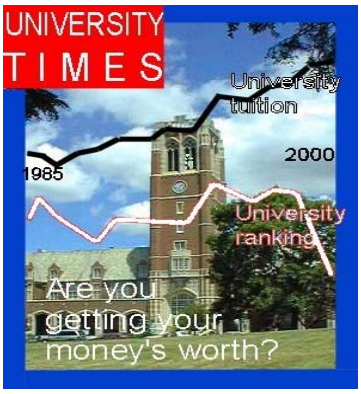



*Figure 1.* Relationship between O-Ring Failures and Temperature

5. In the United States in 1970, there was a draft lottery of 18-year old males. Men were drafted according to the order in which their birthdays were listed. Young men who were born in the last quarter of the year complained that they were discriminated against. Their birthdays appeared disproportionately early in the drawing. Argue their case graphically given the data (see Hollander, 1984).

6. A politician states "Scientists have not yet established any link between smoking and cancer. They have a lot of statistics, but that's all." Note that this politician comes from a region that produces tobacco. a) Discuss this statement and suggest what else this politician might want to have. b) Discuss the nature of any link that might be established.

7. The graphs in Figure 2 were found on the Internet. Critique them. What conclusions do you draw from them? In the one on the left, *Are you getting your money's worth?* The top line plot relates the cost of a certain university's tuition over time, and the bottom line plot displays the university's ranking over the same time period. The graph on the right is a pie chart of regional sales of a product. If the graphs are in error or misleading, draw ones that correct the difficulties.



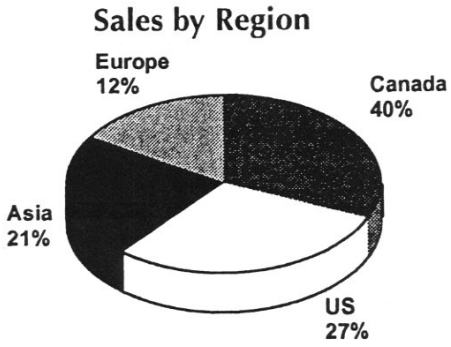




*Figure 2.* Examples of the Graphs Used in the Assessment

8. Refer to the article at www.apa.org/releases/placebo.html. The study indicated that a placebo was found to account for fifty percent of improvement in depressed patients taking antidepressants. So, why bother taking drugs? Discuss the nature of the study regarding the use of a placebo.

9. Research the claim that taking an aspirin a day reduces the chance of suffering a heart attack. Discuss the design of the first experiment done in the late 1980's. Who were the subjects and what were the results? Is the treatment for everyone? (This involves looking at further studies.)

10. It has been argued that the probability of finding at least two people in a class of thirty having the same birthday is (1/366) + (2/366) + (3/366) +…+(29/366) = 435/366 and concluded "This very large ratio means that it is almost certain that two people in any group of thirty will share the same birthday. Not only would you not be surprised to find out this was so, you would expect it." Comment on the computation and on the conclusion. Provide a correct solution if this one is in error.

11. Write a conversation between you and a friend who knows little about statistics why Gallup needs only to interview about 1500 adult Americans to have a fairly good understanding of the proportion of all adult Americans who are in favor of a certain issue.

NOTES ON THE EXAM

What I tried to do in the writing of the examination questions was to summarize the topics that I think everyone should have some conceptual understanding of. The vast majority of the exam purposefully requires answers to be written rather than computations to be calculated. The course covers inferential statistics but only a conceptual knowledge is really needed. Question 1 asks the citizen student to interpret the mode (most Americans consume 20 teaspoons), median (half of all Americans consume less than 20 teaspoons and half, more), midrange (20 is midway between 0 teaspoons of consumption and whatever the maximum is), and the mean (if all Americans were to consume the same number of teaspoons of sugar, it would be 20).

Question 2 is challenging the student to look beyond the *title* of an article, newspaper column, or book. Too often, we take such for granted and especially if we agree with its statement without investigating the basis on which it was made. The statistics techniques in this text are too complicated for all citizens to follow but the point is to have them open the cover and investigate as much as possible. Lott's work is an incredible observational study. Question 3 introduces them to a legal situation in which they may be participants at some point in their lives. I'm interested in the kinds of other information they would want to know before making a decision, e.g., tire pressure, type of surface, insurance claims, etc.

In answering question 4, the point is that all the data were available to the decision makers in charge of the Challenger launch January 28, 1986, but they reasoned not to use all of it. They based their decision only on the seven flights in which there were o-ring failures. The U-shape scatterplot does not provide strong evidence of a connection between flight failure and temperature. However,

when one includes the data from the other sixteen flights in which there were no o-ring failures, it is clear that no flight was successful when the temperature was below 66°F.

An ideal solution to question 5 would show a median plot line of a scatterplot of vertical boxplots with months on the horizontal axis and the birth order of "random" drawings on the vertical. There is a clear drop in the medians of birth orders for the months in the last quarter of the year. I heard this quote in question 6 on the radio but can't find it in print from one of our very interesting and often controversial US senators. I guess the purpose of this question is to get my citizen student to pay attention to what our politicians' understanding of statistics is and the extent to which they use statistics and don't use statistics based on their political aims and positions. Question 7 points out the dangers of omitting description of the vertical axis. The lower graph shows ranking of the university over time; top ranking as at the bottom of the scale and so the impression that this university's tuition is going up but its ranking is going down is erroneous. Regarding three-dimensional piecharts, they should be banned, period! (Compare US and Canada.)

Questions 8 and 9 ask the student to interpret the results of two studies. The first is getting them aware of the place that placebos have played in experiments and to be exposed to a very controversial study regarding them. The aspirin study was first done with middle-aged males and has been extended to other male groups and females since then. The citizen student will have to use the internet on their own to discover some of the recent findings. Question 10 involves probability. Its purpose is to encourage them to challenge what quantitative results they see in print. We read poll results everyday in the newspaper, but tend to believe only those in which we have agreement. Question 11 challenges the citizen student to put into their own words their understanding of inferential estimation.

## CONCLUSION

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 and interpreting data, a necessary knowledge base for them to become able to make critical and informed decisions throughout their lives - as consumers, workers, jurors, citizens. It is hopeful that future generations will be thus equipped through school statistics or through quantitative literacy vis-à-vis Steen's *Mathematics and Democracy* or through innovative mathematics curricula such as Data Driven Mathematics. For the college student, I urge us to review what we are doing and how we are doing it in our college statistics courses so that all of our students will truly use statistics throughout their lives, and will even carry on quantitative conversations at social gatherings! Currently, citizens have at most a "supermarket tabloid" understanding of statistics. To equip them 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, we must consider offering Citizens Statistics 101 courses through community education programs or college/university's continuing education programs.

## REFERENCES

Hollander, M. (1984). *The statistical exorcist -dispelling statistics anxiety* (pp. 121-130). Dekker.

Lott, J. (2000). *More guns, less crime* (2nd edn). University of Chicago Press.

Landwehr, J., & Watkins, A. (1995). *Exploring data* (revised). Dale Seymour Publishers.

Moreno, J. (1998). Statistical literacy: statistics long after school. In *Proceedingds of the 5th International Conference on Teaching Statistic*s (pp. 445-450). International Statistics Institute.

Paulos, J.A. (1988). *Innumeracy: Mathematical illiteracy and its consequences*. NY: Hill & Yang.

Paulos, J.A. (1995). *A mathematician reads the newspaper*. Basic Books.

Steen, L. A. (2001). *Mathematics and democracy: The case for quantitative literacy.* The Woodrow Wilson National Fellowship Foundation.