Numbers in Everyday Life: A Short Course for Adults

Gerald J. Hahn[1], Necip Doganaksoy[2], Ricki Lewis[3],
Jane E. Oppenlander[4], Josef Schmee[5]
[1]GE Global Research, (Retired) Manager-Applied Statistics
[2]GE Global Research, Principal Technologist-Statistician
[3]Geneticist and Science Writer
[4]Adjunct Professor, Union Graduate College
[5]Professor Emeritus, Union Graduate College

Abstract
We recently developed and gave a five-session course on *Numbers in Everyday Life* at the Union College Academy for Lifelong Learning. We describe how we put the course together, what we presented, and how things worked out. Copies of most of our presentation slides and other materials can be accessed via a cited Web site.

Key Words: Communications, introductory statistics course, statistical literacy, statistics for consumers, teaching statistics.

1. Introduction

Statistical literacy has been defined by Katherine Wallman as "the ability to understand and critically evaluate statistical results that permeate our daily lives." On a more technical level, Schield (2010) defines "Statistical literacy is the ability to read and interpret summary statistics in the everyday media: in graphs, tables, statements, surveys and studies." Building a statistically literate community is one of our profession's major responsibilities and challenges, as emphasized by numerous ASA presidents, including Bob Mason, Richard Scheaffer, and Katherine Wallman. David Moore, another ASA president, distinguishes statistical literacy—"what every educated person should know about statistical thinking"—from statistical competence—"roughly the content of a first course for those who must deal with data in their work." The need to build statistical literacy is also reflected in the ASA strategic plan, with an objective to "stimulate public awareness of the role of statistics and statisticians in issues that affect public life."

Many, including Debora J. Rumsey (2002) in her *Journal of Statistics Education* article, "Statistical Literacy as a Goal for Introductory Statistics Courses," have urged that statistical literacy concepts be included in introductory statistics courses. Recent texts have worked toward this, and college courses devoted to statistical literacy have been developed.

In recent years much emphasis has rightfully been placed on statistics in K to 12 education. The Advanced Placement course in statistics now attracts over 100,000 high school seniors yearly. Also, many college students today take an introductory course in statistics. Such courses tend to teach students methods for analyzing data. However, they frequently fail to focus on the use (and abuse) of statistics in *everyday* life.

2. Addressing the Need

Various schools now offer a statistical literacy course, sometimes directed at "adults" as well as full-time students. Such a course does not appear to be available in our area (Schenectady, New York), and even if it were, many might regard a full-fledged college course somewhat over-bearing and requiring a greater commitment of time and effort than they are willing to make.

Short courses offered by adult education programs offer—such as the locally based Union College Academy for Lifelong Learning (UCALL)--provide an excellent alternative venue for promoting statistical literacy. It was our good fortune to be invited by UCALL to develop and offer (in the spring of 2008) such a course.

As its name suggests, UCALL is a "program for adults who are committed to lifelong learning. Membership is open to any adult who wishes to continue learning in an intellectually stimulating environment. Courses cover a wide range of topics…Classes are held for five 2-hour sessions in the fall and spring." The courses are offered in the late morning and early afternoon and, therefore, attract principally intellectually curious retirees. The spring 2008 curriculum, in addition to ours, included courses on:

- The Roaring Twenties
- Leonardo daVinci
- Behind the Scenes of Local Theater Productions
- Eastern Religions—Their Role in Civilization's Great Spiritual Transitions
- Four Operas from the Beginning of the 20th Century (also taught by one of us, Josef Schmee; who said that statisticians were one-dimensional?)

2007/8 season membership in UCALL cost $65, which permitted participation in any of the 12 courses, with an additional charge of $25 for each of the first two courses per semester and no charge for additional courses.

3. Course Organization and Start-Up

We met with our course coordinator about nine months prior to the course (coincidentally, the time we understand it takes to have a baby) to lay the foundation and make preliminary plans.

Sizing up our audience. We recognized that members of our audience had some important things in common. All had appreciable life experience, a high level of intellectual curiosity, and sufficient interest in the subject to sign up for such a course.

There were, however, also some appreciable differences. Some participants had likely never taken a statistics course; others, in addition to having training in statistics, might have used it in their work. It turned out that the class registrants included a one-time supermarket manager, a librarian, a social worker, a GE research biologist, and a psychology professor—to name just a few. We, therefore, needed to strike a balance between confusing some of the participants with too advanced material and boring others by too elementary presentations.

Scoping the course. One thing that most participants did *not* want was a mini-course in statistical methods and theory. Few were likely to perform statistical analyses themselves in the future. But all are exposed daily to numbers in the media and on the

Internet—and wanted to understand better what these really meant. In short, we were dealing with consumers--not producers--of statistics.

We therefore needed to focus the course around the use and abuse of statistics in specific application areas and introduce technical concepts only when absolutely required to make the application meaningful. The most advanced concept that we would talk about would be the idea (but not the specific method of construction) of a confidence interval.

Course topics. We then began to lay out general topics for each of the five classes. The first class (led by Gerry Hahn and Jane Oppenlander) would lay the foundation via a general introduction of the topic, a few thought-provoking examples, and a discussion of some key concepts.

We then selected application areas that we felt would be of greatest interest to our audience. We decided to devote the middle three classes to:

- Public opinion polls and election forecasts—a subject of particularly strong interest in 2008, a presidential election year (led by Josef Schmee),
- Health studies (led by Ricki Lewis, a science writer and scientist by profession, and the only one of us who was not a statistician),
- Business and industrial applications (led by Necip Doganaksoy).

The fifth and final class (led by Gerry Hahn and Jane Oppenlander) would be structured similar to the first class, with additional examples and pointers and a course wrap-up.

Some other preliminaries. Even though the intent of the course was clear, we did not have a formal title. One of us suggested Statistics for the Consumer. Our course co-coordinator gently advised us that the use of the word statistics—even in the context suggested—could be a definite turn-off. We reluctantly swallowed our professional prides and settled on Numbers in Everyday Life.

We then searched for similar courses offered elsewhere to guide us. We found only one—a course, at the time under development, taught by Dick DeVeaux at the Berkshire Institute for Lifelong Learning, focusing on data mining.

Next, we developed a course description (Figure 1) that succinctly summarized the course for inclusion in the UCALL 2008 spring semester brochure.

> Open a newspaper or turn on a TV—numbers are everywhere from political polls to health studies to sports. They can provide valuable, even life-and-death information, or mislead. This course will provide insights to interpreting numbers and being well-informed citizens. Seasoned experts (all Ph.D.'s) will help you understand the latest statistics from medical studies, public opinion polls, business and industry, to this week's media and more.

Figure 1: Course description: Numbers in everyday life

Class preparation. Each of us then prepared our class material, communicating with each other by email to avoid overlap and to make sure that all the major concepts that we wished to convey were covered (some more than once in different contexts). We met

once more, a week before the start of the course, to help ensure that the course was well integrated and to exchange last-minute ideas.

Registration. We fretted about whether anybody would actually be interested in taking such a course. We were pleasantly surprised to find that more than 60 people had signed up—not quite as good as Josef's (long-running) opera class, but still a very respectable showing. Our choice of course title, indeed, seemed to have been a wise one!

We used the UCALL facilities to send a welcome message to the registrants, providing further course detail and inviting students "to watch for examples of media items dealing with numbers that you would like to see discussed and to bring these to class or email them to one of us." We also noted that perhaps, not coincidentally, the course had been scheduled to start only two days after April Fools' Day.

4. Subject Specifics

So what did we come up with? In this section we summarize the material that we chose to include in the course.

4.1 Class 1: Some Examples and Basic Concepts

This session focused on the following topics:

Course goal and overview: We reviewed the major goals of the course; these were to provide insights into numbers and demonstrate their value, and to highlight abuses. We also stressed that the course would *not* cover:

- Accounting numbers (and investment advice)
- Government statistics (not enough time)
- Statistical methods per se

Studies of impact of NY State Clean Indoor Act and of marriage on longevity: We examined the assertion that "fewer New Yorkers have been treated for heart attacks since the State's wide ranging no smoking law took effect four years ago" and studies linking marriage to longevity. These, and some further examples, allowed us to talk about the common confusion between correlation and causation (a subject to be reiterated throughout the course) and to differentiate between observational and controlled studies.

Study of impact of prayer on outcome of bypass surgery: This example was triggered by a recent local newspaper headline (incorrectly) asserting that "Study finds prayer may make patients worse." We discussed some of the nitty-gritties of this controlled study and, along the way, introduced the concepts of relative risk and of a 95% confidence interval. Some basic concepts illustrated by this example were:

- Read beyond the headline
- Find out how study was conducted and its limitations
- Publication in refereed journal adds credibility

That magical bell-shaped curve: A discussion of when one might expect observations to follow the "so-called normal distribution" and when not. Class participation in the time-tested birthday example effectively illustrated the concept: constructing a histogram of class members' *individual* birthdates within a month and those of their two closest relatives (i.e., three numbers ranging from 1 to 31 for each participant) and then

constructing a similar histogram for the *averages* of these three numbers. This demonstration of the central limit theorem (only casually identified as such) led to class discussion of common phenomena that would and would not be expected to be normally distributed.

Beware of "on the average": We differentiated between the median and the mean with examples dealing with family income and President Bush's 2005 claim that "on the average the folks who sign up for the (then new) prescription drug program are going to save $1,300 per year."

What is data mining?: We illustrated the basic concepts of data mining using the Advanced Scout data mining software. We discussed how coaches in the National Basketball Association use this program to formulate game strategies and to assess the effectiveness of the resulting decisions.

4.2 Class 2: Public Opinion Polls and Election Forecasts

This session was organized around five topics:

Why we need polls: We looked at the need in a democracy to track the priorities and opinions of voters and how these differ from those of pundits and politicians. We showed examples of short-term and long-term changes in voters' opinions.

What is a poll?: We described a poll as a process from constructing a questionnaire to communicating the results. We gave examples of properly and improperly (e.g., leading, double barrel) phrased questions. We also discussed personal interviews, telephone interviews, as well as mailed and on-line questionnaires. We described the Literary Digest poll of 1936 and how its deficiencies led to the modern poll.

How polls are conducted: We differentiated between the sampling frame and the population and problems resulting from this difference. We discussed various selection methods, from random digital dialing to Internet opt-in surveys to non-probability methods. We gave a simple example to show how weighting of subpopulations affects overall estimates. An example from the 2008 primaries, for which two polls conducted by the same polling organization at the same time disagreed, demonstrated the difficulty in conducting polls.

Why polls work: We defined and illustrated statistical margin of error and discussed the relationship of polling precision and accuracy to sample size, non-response rate, how questions were asked, and the timing of the poll, especially in volatile situations. The 2008 Democratic primaries of New Hampshire and California provided opportunities to discuss reasons for the divergence of results amongst various pollsters and also how these compared with the eventual primary results. We mentioned margin of error of subgroups and problems currently facing pollsters resulting from changes from land-based to cellular telephones.

When to trust a poll: We summarized problems inherent in pre-election polls and gave some guidelines on what constitutes a good poll. We used a New York Times sidebar, "How the Poll Was Conducted," to review the key components of a well-constructed poll. We cautioned that polls, no matter how well executed, may produce results that do not match the eventual outcome. We ended with a tongue-in-cheek prediction of the

outcome of the 2008 presidential race to show the limitations of polls seven months before an election. We also referred the class to several relevant websites.

4.3 Class 3: Health Studies

The session instructor examined sources of science news and feature stories. Slides summarized concepts and liberally used cartoons, illustrations, and photographs. This session consisted of many examples loosely organized into three sections.

Anatomy of a medical journal: To get a flavor of how medical journals use and present statistics, we examined a recent issue of The New England Journal of Medicine (2/28/08). This is important because the media often report information from such journals. The eight articles in this issue illustrated such concepts as the value of negative evidence; the importance of context, bias, and invalid assumptions; confusing correlation with causation; odds ratio to evaluate cancer risk; sample size requirements; and cost/benefit analysis in utilizing an expensive medical scanning technology. Some examples did not fit a specific category but were still intriguing, such as the effect on incidence of cystic fibrosis before and after prenatal screening (enabling women to terminate affected pregnancies) became common.

The drug approval process: A short bridging discussion considered the steps of the drug approval process in the US, from basic research, to preclinical (animal) studies, to clinical trial design. We talked about how health studies enter the news, and how the media misuse terms such as "proof" and "theory." A handout listed commonly used terms, and covered incidence and prevalence; risk; association, correlation, and causation ("do firefighters cause fires?"); the importance of determining the underlying mechanism (smoking and lung cancer); and future additions to the drug approval process (biomarkers and pharmacogenomics).

Sources of distortion: We used news releases (mostly from www.eurekalert.org) to illustrate and categorize statistical errors and misinterpretations. The examples dealt with inappropriate extrapolations; errors of omission and study drop-outs; test duration and sample size limitations (the reason why many blockbuster drugs show adverse effects only *after* they have been marketed); meta-analyses; confounding factors; and flawed surrogate markers.

The session concluded with take-home messages in evaluating health studies in the news. Students should be on the lookout for the underlying logic; missing or irrelevant information; confounding factors; lack of controls; and whether researchers or reporters have an agenda or conflict.

4.4 Class 4: Business and Industrial Applications

This session was organized in two parts.

The first part illustrated common pitfalls and misuses of statistics in business and industry, and their consequences. It provided the following recommendations:

Find out how numbers are defined: This discussion focused on the Mars Climate Orbiter which perished during entry into Mars's orbit. A contract engineering team had used English units instead of the metric system specified by NASA.

Graphical displays are highly useful, but can also mislead: A sample dataset displayed graphically created noticeably different impressions by simply adjusting the scales of the graph. This led to a lively class discussion about the potential pitfalls in graphs used to display financial data (e.g., stock prices, company performance).

Be wary of advocates with numbers: In ad campaigns run in parallel, various cell phone companies claimed to be the "most reliable" service provider. The discussion highlighted the importance of understanding the study design and the definition of terminology.

Find out how numbers were obtained: The selective use of data on O-ring failures, resulting in the disastrous decision to launch the NASA Space Shuttle Challenger on an unusually cold Florida morning in January 1986, was discussed.

The second part provided examples of how statistics is used effectively in business and industry:

Identifying differences and seeking their causes: A comparison of key performance characteristics at two plants showed that lots manufactured at Plant A were consistently closer to their target value than those manufactured at Plant B. Further examination revealed that Plant A performed additional undesirable product processing.

Improving operations: A television station wanted to improve the quality of its closed captions (i.e., subtitles for the hard-of-hearing). Appropriate data were gathered and analyzed on the quality (accuracy, completeness, timeliness, readability, etc.) of the station's past captions and those of competing stations for varying program types (e.g., news, sporting events, reality shows). These results were used to identify problem areas and potential improvements.

Preventing fraudulent activities: This study dealt with the effectiveness of sending warning letters to Internet users that were illegally sharing copyrighted files (e.g., music, videos, software, books). The discussion highlighted the key features of observational studies versus randomized trials. The available data helped structure a further study, the results of which were used to develop a strategy for reducing future fraudulent activities.

4.5 Class 5: Further Examples and Wrap-Up

The final session focused on the following topics:

Quantifying the subjective—college rankings: An examination of subjective evaluations, such as the U.S. News and World Report yearly college rankings, and their construction.

Testing in schools: A description of how New York State currently assesses a school's conformance to the requirements of the 2001 No Child Left Behind Act and its

limitations, and an introduction to the more statistically refined "added value assessment." Not unexpectedly, this example generated much discussion.

More on data mining: We examined:

- How Wal-Mart uses data mining to help its bottom line. This included a discussion of the company's massive data warehouse and how it is used to discover customer buying patterns, manage suppliers and inventory, and evaluate store and employee performance.
- The issues surrounding the use of databases and data mining and personal privacy. This included examples regarding the sale of firearms, homeland security, and the use of retail loyalty programs.
- A new project using data mining and complex systems modeling to assess pollution in the Lake Champlain watershed.

Diagnosis of a "one-in-a million" chance event: An assessment of a media claim that "in the 2008 Syracuse (NY) Democratic primary Clinton and Obama each got 6,001 votes…The odds of that happening are less than one in a million." This example showed how impressive and surprising results make news—but just because something appears in the newspaper doesn't make it right.

Sports applications: Overviews from baseball, basketball and football and examination of the question "what are the chances that a 0.300 hitter with a 0.350 average in his first 100 at bats will end the season with a 0.350 average?"

Good and bad graphics: This focused on a compilation of misleading graphics taken from Best (2001 and 2004).

Number studies further knowledge: To end on a positive note, we summarized the benefits of numbers studies by a variety of examples. These included resolving the disputed authorship of the Federalist papers, identifying archeological finds, assessing gender and race discrimination, evaluating global warming, and an investigation to determine whether birds can differentiate jet engine noise from mating calls.

Some good reading and surfing: We provided a categorized handout of further reading and computer surfing opportunities, highlighting

- The books by Best (2001 and 2004), Peck et al (2006) and Utts (1999) and Chance Magazine.
- The Chance News (http://www.dartmouth.edu/~chance/) and Carl Bialik (the Wall Street Journal Numbers' Guy (http://blogs.wsj.com/numbersguy/) Web sites.

We also responded to a request to comment on The Black Swan (Taleb, 2007) and ended with what we regarded as the major "take-aways" that we had tried to emphasize in the course (see Figure 2).

5. Some Observations

In responses to the end-of-class survey, the great preponderance of participants felt that the course had increased their knowledge of the subject and that they had enjoyed coming

to class. (However, we did wonder how well the views of the 50% class members who responded to the survey reflected those of the entire class)

We were a little disappointed to receive essentially no suggestions on how the class could have been improved and no responses to our request for participants to bring in their own examples from the media. But we were pleased with the extensive class discussion that each of the topics generated and that we tried to encourage. Two participants did comment to us privately that, because of the sensitive nature of the subject, we should not have included the example dealing with the impact of prayer on the outcome of bypass surgery.

- Numbers are an essential and highly valuable element of numerous human endeavors—you can't escape them
- Always ask
  - Who is taking/reporting the numbers?
  - How were they obtained?
  - Have they been peer-reviewed?
  - What are the underlying assumptions?
- Be wary of
  - Advocates' numbers
  - Cherry-picking
  - Before and after comparisons
- Remember news media seek newsy/surprising numbers
- Appreciate limitations of observational studies and differentiate correlation from cause and effect
- Gold standard is controlled (randomized) experimentation—but often not attainable
- Recognize uncertainty: Nothing is certain, but death and taxes (Benjamin Franklin)
- Let numbers help you gain understanding--not intimidate you!

CLASS MOTTO: Numbers are highly useful, but can be readily abused—handle with care!

Figure 1: Course take-aways

We also felt that we succeeded in limiting the technical discussion to the essentials that the participants needed to know. As mentioned, the most complex idea presented was that of a confidence interval. Not unexpectedly, some participants found this concept difficult to grasp. When asked how exactly one goes about constructing such intervals, we copped out by saying that "this is what is taught in introductory statistics courses."

As indicated, our course—in addition to being more oriented towards consumers of statistics than most--attracted the other end of the age spectrum from that of most existing offerings. A similar offering might, however, also serve the needs of those in the middle.

We note that a mini-course, such as the one we offered, might draw a larger audience than a full-fledged course in that it requires less of a commitment of time. And it makes it even more important that the limited available time be used in an optimum manner with the goal of building statistical literacy. But even a mini-course will likely attract only a

fraction of the population. So a remaining question is "how can we build statistical literacy among those that are not willing to take a formal course?"

Finally, we observe, that doing this undertaking as a "gang of five" worked out extremely well. It is questionable whether any one of us individually would have had the time or inclination to put such a course together. Moreover, it exposed the class participants to a variety of perspectives and allowed each of us to focus on the topics that we knew most about. Hearing basically the same concept repeated by various instructors in different contexts (e.g., correlation versus causation) also reinforced important points. And, not unimportantly, it was fun working together.

6. Availability of Class Materials

This paper is an extended and updated version of Hahn, G. et al (2009). Copies of the following materials can be obtained from Milo Schield's web site (statlit.org):

- Slides from our presentation at the 2010 Joint Statistical Meetings.
- Slides of 4 of the 5 UCALL course presentations.
- A detailed summary of the fifth presentation (on health studies).
- Class summary and reading list.
- Copy of Hahn, G. et al (2009).

To access these materials on statlit.org

- Go to Statlit News
- Go to Statlit 2009
- Go to Numbers in Everyday Life
- Hit links shown about third of the way down (mostly on the right hand side)

For further information or questions, contact gerryhahn@yahoo.com or doganaksoy@research.ge.com

Acknowledgements

We thank Jim Comly for inviting us--on behalf of UCALL--to teach this course, for serving as course coordinator, and for his many useful inputs. We also thank Valerie D'Amario and Nancy Mancini of the UCALL office for their valuable support.

References

Best, J. (2001), *Dammed Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*, University of California Press.

Best, J. (2004), *More Dammed Lies and Statistics: How Numbers Confuse Public Issues*, University of California Press.

Hahn, G.J. et al (2009), Numbers in Everyday Life: A Short Course for Adults, *Amstat News*, #380, 16-19.

Peck, R et al (2006), *Statistics: A Guide to the Unknown*, 4th edition, Duxbury Press.

Rumsey, D. J. (2002), Statistical Literacy as a Goal for Introductory Statistics Courses, *Journal of Statistics Education*, Vol. 10, Number 3.

Schield, M. (2010), Assessing Statistical Literacy: Take CARE, Chapter 11 in *Assessment Methods in Statistical Education*, Wiley.

Taleb, N. N., (2007), *The Black Swan, the Impact of the Highly Improbable,* Random House.

Utts, J. M. (2004), *Seeing through Statistics*, 3rd edition, Duxbury Press.