

Teaching Bits: A Resource for Teachers of Statistics

Journal of Statistics Education v.8, n.3 (2000)

Robert C. delMas
General College
University of Minnesota
354 Appleby Hall
Minneapolis, MN 55455
612-625-2076

delma001@tc.umn.edu

William P. Peterson
Department of Mathematics and Computer Science
Middlebury College
Middlebury, VT 05753-6145
802-443-5417

wpeterson@middlebury.edu

This column features "bits" of information sampled from a variety of sources that may be of interest to teachers of statistics. Bob abstracts information from the literature on teaching and learning statistics, while Bill summarizes articles from the news and other media that may be used with students to provoke discussions or serve as a basis for classroom activities or student projects. Bill's contributions are derived from Chance News (http://www.dartmouth.edu/~chance/chance_news/news.html). Like Chance News, Bill's contributions are freely redistributable under the terms of the GNU General Public License (<http://gnu.via.ecp.fr/copyleft/gpl.html>), as published by the Free Software Foundation. We realize that due to limitations in the literature we have access to and time to review, we may overlook some potential articles for this column, and therefore encourage you to send us your reviews and suggestions for abstracts.

From the Literature on Teaching and Learning Statistics

The College Mathematics Journal

"The Super Bowl Theory: Fourth and Long"

by Paul M. Sommers (2000). *The College Mathematics Journal*, 31(3), 189-192.

The author notes that the Super Bowl Theory (that there is a reliable correlation between stock market performance and Super Bowl results) was surprisingly accurate and reliable in the early years of the Super Bowl. The theory has produced unreliable predictions in recent years. The article presents a dataset that contains Super

Bowl and Dow Jones data from 1967 to 1998, describes a regression model based on the theory, and provides an example of how to assess the explanatory power of the theory.

" t -Probabilities as Finite Sums"

by Neil Eklund (2000). *The College Mathematics Journal*, 31(3), 217-218.

The author illustrates how the expression for the t -probability formula can be written as a finite sum. The finite-sum expression can be used to write a program for a programmable calculator to accurately determine the probability for a given t -value and degrees of freedom.

"Food and Drug Interaction: What Role Does Statistics Play?"

by Tom Bradstreet (2000). *The College Mathematics Journal*, 31(4), 268-273.

The author describes and provides data from a study that was conducted to investigate the magnitude of the food interaction of a new drug. An example of how the dataset can be used to illustrate normal theory ANOVA with subsequent confidence intervals is provided. The author states that the dataset can also be used to illustrate the ideas of paired vs. independent data, hypothesis testing procedures for the paired t -test, the Wilcoxon signed rank test and the corresponding Hodges-Lehmann estimator and Moses confidence interval, outliers, exploratory data analysis methodologies, and the use of graphical displays. A URL is provided at the end of the article for access to twenty-two other datasets.

Journal of Educational and Behavioral Statistics: Teacher's Corner

"Note on the Odds Ratio and the Probability Ratio"

by Michael P. Cohen (2000). *Journal of Educational and Behavioral Statistics*, 25(2), 249-252.

The note does not provide an explanation of the odds ratio, although the author provides useful references. The author provides examples that can be used in the classroom to help students learn the similarities and differences between the odds ratio and the probability ratio (relative risk), and to help students to start thinking in terms of odds ratios.

Journal for Research in Mathematical Education

"On the Complexity of Schools in Contemporary Society: How Can Research Inform Us About Mathematics Learning and Teaching?"

by the NCTM Research Advisory Committee (P. Tinto, C. Fernandez, J. Garfield, D. Grouws, G. Lester, G. Martin, D. Majee-Ullah, and G. Stimpson) (2000). *Journal for Research in Mathematical Education*, 31(5), 520-523.

This report gives brief updates on plans to publish "A Research Companion to the NCTM Standards," the formation of a Standards Impact Research Group (SIRG), the report completed by the Task Force on Mathematics Teaching and Learning in Poor Communities, and plans for the publication of a "Research Briefs" series that will address questions about mathematics curriculum, teaching, and learning posed by teachers,

teacher educators, administrators, and parents. The second half of the report outlines five challenges for the mathematics education research community.

"Making Sense of the Total of Two Dice"

by Dave Pratt (2000). *Journal for Research in Mathematical Education*, 31(5), 602-625.

Abstract: Many studies have shown that the strategies used in making judgments of chance are subject to systematic bias. Concerning chance and randomness, little is known about the relationship between the external structuring resources, made available for example in a pedagogic environment, and the construction of new internal resources. In this study I used a novel approach in which young children articulated their meaning for chance through their attempts to "mend" possibly broken computer-based stochastic gadgets. I describe the interplay between informal intuitions and computer-based resources as the children constructed new internal resources for making sense of the total of 2 spinners and 2 dice.

The American Statistician: Teacher's Corner

"Applying Cognitive Theory to Statistics Instruction"

by Marsha C. Lovett and Joel B. Greenhouse (2000). *The American Statistician*, 54(3), 196-206.

Abstract: This article presents five principles of learning, derived from cognitive theory and supported by empirical results in cognitive psychology. To bridge the gap between theory and practice, each of these principles is transformed into a practical guideline and exemplified in a real teaching context. It is argued that this approach of putting cognitive theory into practice can offer several benefits to statistics education: a means for explaining and understanding why reform efforts work; a set of guidelines that can help instructors make well-informed design decisions when implementing these reforms; and a framework for generating new and effective instructional innovations.

"Collections of Simple Effects and Their Relationship to Main Effects and Interactions in Factorials"

by Oliver Schabenberger, Timothy G. Gregoire, and Fanzhi Kong (2000). *The American Statistician*, 54(3), 210-214.

Abstract: When analyzing data from a cross-classified experiment, one of the primary interests lies in estimating and testing main effects and interactions. Frequently, one is also interested in comparing the levels of one factor at a given level of another factor, particularly if interactions are present. Such collections of simple main effects are termed *slices* herein. Based on unfolded designs in which effects are represented by generic sets of orthogonal contrasts among cell means, the relationships between the contrast sets defining main effects, interactions, and slices in terms of Kronecker representations and projection properties are examined. The material is appropriate in a first course on linear models and/or a course in experimental design with linear algebra prerequisite to demonstrate the relationship and interpretation of various effects in a factorial setting. An example from production quality control is presented.

"The Mighty Bonus Point"

by George D. Kesling (2000). *The American Statistician*, 54(3), 215-216.

Abstract: This article suggests a small change in grading that can cause a big change in learning í- "the mighty bonus point." Bonus points that contribute no more than two percent of a student's grade can help create a group atmosphere where everyone helps the group. Bonus points are awarded for outstanding contributions to the class and are used to encourage peer tutoring and general group learning.

Teaching Statistics

A regular component of the Teaching Bits Department is a list of articles from *Teaching Statistics*, an international journal based in England. Brief summaries of the articles are included. In addition to these articles, *Teaching Statistics* features several regular departments that may be of interest, including Computing Corner, Curriculum Matters, Data Bank, Historical Perspective, Practical Activities, Problem Page, Project Parade, Research Report, Book Reviews, and News and Notes.

The Circulation Manager of *Teaching Statistics* is Peter Holmes, Peter.Holmes@ntu.ac.uk, RSS Centre for Statistical Education, University of Nottingham, Nottingham NG7 2RD, England. *Teaching Statistics* has a web site at <http://science.ntu.ac.uk/rsscse/ts/>.

Teaching Statistics, Autumn 2000 Volume 22, Number 3

"CensusAtSchool 2000" by Doreen Connor, Neville Davies, and Peter Holmes, 66-70.

Abstract: The article covers the development of an ambitious Internet-based project to conduct a simple census of the schoolchildren of England and Wales linked to the 2001 UK National Census.

"Do the Lengths of Pebbles on Hastings Beach Follow a Normal Distribution?" by Anne Brooks, 73-76.

The author describes a research activity conducted with secondary school students. The task involved the collection of some 600 pebbles from a beach in order to estimate the distribution of the pebble lengths. Students learned about the characteristics of a normal distribution within a real-life context. In addition, the activity provided experience in dealing with large datasets, generating graphs, calculating statistics, identification and use of appropriate data collection and analysis techniques, interpretation of results, statistical decision-making, and communication of findings.

"The Case Against Histograms" by David L. Farnsworth, 81-85.

The author describes a homework assignment that illustrates to students how decisions about the number of intervals and interval widths to use when creating histograms can influence inferences about the shape of a distribution.

"Learning Statistics on the Web í- DISCUSS" by Neville Hunt and Sidney Tyrrell, 85-90.

Abstract: This article discusses the use of the Internet in teaching and learning statistics and describes how interactive spreadsheets can be integrated with Web-based resources.

"On Simulation and the Teaching of Statistics" by Ted Hodgson and Maurice Burke, 90-96.

Abstract: The use of simulation as an instructional tool can promote a deep conceptual understanding of statistics and lead to misunderstandings. Teachers need to be aware of the misconceptions that can arise as a result of

simulation and carefully structure classroom activities so as to derive the benefits of this powerful instructional tool.

Topics for Discussion from Current Newspapers and Journals

"Study Disputes Success of the Brady Law"

by Fox Butterfield, *The New York Times*, 2 August 2000, p. 12.

"Two Economists Give Far Higher Cost of Gun Violence"

by Fox Butterfield, *The New York Times*, 15 September 2000, p. 12.

The first article is based on a study in the *Journal of the American Medical Association* ("Homicide and Suicide Rates Associated With Implementation of the Brady Handgun Violence Prevention Act" by Jens Ludwig and Philip J. Cook, *JAMA*, 2 August 2000, Vol. 284, No. 5, pp. 585-591). Gun homicides have declined overall since the Brady law was passed. However, 18 states already had background checks and waiting periods before the law was enacted, while 32 states enacted such restrictions in response to the law. If the law were really having a positive effect, one might expect the latter group to show a faster rate of decline in gun homicide. But the study did not find such an effect.

The Clinton administration joined Brady law supporters in criticizing the study. They point out that the study did not explicitly consider the possible effects that the Brady law may have had on the 18 "control" states. Specifically, as the authors of the study themselves acknowledged, "...it is possible that the Brady Act may have had a negative association with homicide rates in both the treatment and control states by reducing the flow of guns from treatment-state gun dealers into secondary gun markets." The secondary gun markets referred to here include gun shows and private sales, which are not subject to the Brady law. Such transactions account for an estimated 30 to 40% of the gun market.

The second article describes Ludwig and Cook's recently published book *Gun Violence: The Real Costs* (Oxford University Press). The central thesis of the book is that the real costs of gun violence far exceed the medical costs and lost workplace productivity of gunshot victims. The authors' total cost estimate of \$100 billion a year is based on a technique that economists call "contingent valuation," which assesses how much it is worth to people to avoid a problem. The authors conducted a telephone survey of 1200 people, asking them how much they would pay annually to reduce criminal gunshot injuries. This led to an estimate of \$80 billion nationally. The remaining \$20 billion was derived from previous analyses of lost productivity and the use of jury awards to place a value on a lost life.

Critics of the book note that Ludwig and Cook's accounting fails to include any benefits of guns. They also question the highly subjective element of putting a dollar value on a human life.

"Behind the Good News on Cancer There's the Same Old Bad News"

by H. Gilbert Welch, *The Los Angeles Times*, 19 July 2000, p. 9.

What is the right measure of progress in the war on cancer? One widely quoted statistic is the five-year survival rate, defined as the proportion of people alive five years after being diagnosed with the disease. Welch was a co-author of an article in the *Journal of the American Medical Association* that questioned the usefulness of this statistic (Welch, Schwartz, and Woloshin, "Are Increasing 5-Year Survival Rates Evidence of Success Against Cancer?" *JAMA*, 14 June 2000, Vol. 283, No. 22, pp. 2965-2978). In the *Times* article, he presents two simple examples to illustrate that the survival rate can increase even when there has been no change in the effectiveness of treatment.

The first way this can happen occurs when new tests for a particular cancer permit earlier detection of the disease. Welch asks us to imagine a group of men who died of prostate cancer at age 78, the current median age of death for men with this disease. If these men had been diagnosed to have prostate cancer at age 75, their five-year survival rate would be 0. But if they had been diagnosed with the disease at age 70, this rate would be 100%. In other words, early detection of a disease can cause the five-year survival rate to increase even with no change in the way the disease is treated.

A second problem with using survival rate is the fact that some cancers have both progressive and non-progressive forms. In the progressive form, the disease can progress to the point where it is life-threatening. In the non-progressive form, there will typically be no symptoms, and the disease will not result in death. Prostate cancer is an example. Before the recent trend towards aggressive testing, virtually all patients diagnosed with prostate cancer would have exhibited symptoms of the progressive form. Thus if 1000 men were diagnosed with prostate cancer in 1950 and 400 of them were alive five years later, the five-year survival rate would be 40%. But with the new diagnostic tests for prostate cancer, if 1000 men are found to have prostate cancer, some of these will have the non-progressive form and thus not be at risk of death from the disease. As before, this will increase the observed survival rate even if there is no change in the treatment of the disease.

The original *JAMA* article compares data for the five-year survival rate, the mortality rate, and the incidence rate for 20 forms of cancer for 1950-1954 and 1989-1995. The results are presented in the following table:

Cancer	5-year survival, %		Absolute increase in 5-year survival, %	% change (1950-1996)	
	'50-'54	'89-'95		Mortality	Incidence
Prostate	43	93	50	10	190
Melanoma	49	88	39	161	453
Testis	57	96	39	-73	106
Bladder	53	82	29	-35	51
Kidney	34	61	27	37	126
Breast	60	86	26	-8	55
Colon	41	62	21	-21	12
Rectum	40	60	20	-67	-27
Ovary	30	50	20	-2	3
Thyroid	80	95	15	-48	142
Larynx	52	66	14	-14	38
Uterus	72	86	14	-67	0

Cervix	59	71	12	-76	-79
Oral Cavity	46	56	10	-37	-38
Esophagus	4	13	9	22	-8
Brain	21	30	9	45	68
Lung	6	14	8	259	249
Stomach	12	19	7	-80	-78
Liver	1	6	5	34	140
Pancreas	1	4	3	16	9

Observe how the trends in the five-year survival rates are all very encouraging. Nevertheless, in terms of mortality rates, the picture is a disturbing mix.

"TV Advertising Drives Fight Over Size of Spanish Audience"

by Jayson Blair, *The New York Times*, 17 July 2000, A1.

Univision, the largest Spanish-language television broadcaster in the US, has charged that Spanish-speaking households are underrepresented in the sample used for Nielsen ratings in the New York area. There are 6.8 million households with televisions in this area, about 1 million of which are estimated to be Hispanic. Nielsen has viewing meters in a total of 500 households. Although Nielsen believes its sample includes a reasonable fraction of Hispanic households, officials acknowledge that their ratings do not distinguish households that speak primarily Spanish.

To further investigate the language issue, Nielsen conducted a door-to-door survey to estimate the size of the primarily Spanish-speaking population. Their results indicated that 43 of the 500 viewing-meters should be placed in "Spanish-dominant" households, whereas currently there are only 21. Needless to say, English-language broadcasters were not pleased with the prospect of losing meters from English-speaking households. They criticized the methodology of the language survey, because interviewers initiated conversations in Spanish and allegedly asked leading questions. Nielsen agrees that more research should have been done into "Spanish-language polling techniques," but still feels that Spanish-speaking households are being undercounted. Nevertheless, concerns about the survey have convinced Nielsen to put on hold their plans to change the meter locations.

Apart from the survey, there is reason to believe that Univision's concerns may be justified. The article reports that since 1995 Nielsen has maintained a separate rating called the Nielsen Hispanic index (NHI), which it believes gives a more accurate account of Hispanic viewership. Nielsen's primary rating service is called the Nielsen station index (NSI). Reproduced below is a comparison from a sidebar to the article. It shows Nielsen NSI figures for the number of viewers aged 18-49 for the major evening news broadcasts. The figures in parentheses are NHI estimates for Univision.

	February weekdays	May "sweeps"
Univision	137,200 (289,600)	165,100 (217,600)
ABC	228,800	289,600

CBS	108,600	116,700
NBC	289,200	230,200

"Vegetarians Are More Likely to Have Daughters"

by Helen Rumblelow, *The Times* (London), 7 August 2000.

"Plant Chemicals Could Play Small Part in Deciding Gender"

by Dr. Thomas Stuttford, *The Times* (London), 7 August 2000.

Both articles concern a 1998 study of nearly 6000 pregnant women in Nottingham, England, 5% of whom did not eat meat or fish. Among the vegetarians, the ratio of boy babies to girl babies was 85 to 100, while for the meat and fish eaters, the ratio was 106 to 100. According to the articles, the second ratio reflects the overall national average for Great Britain.

The only other study on how diet affects the sex of offspring showed that high levels of magnesium, potassium, and calcium result in more boys. However, there is no evidence that vegetarian diets are low in these elements. The new study also found that vegetarians are less likely to smoke during pregnancy than non-vegetarians (10% versus 25%). Previous research had indicated that nonsmokers tend to produce more boys. One of the researchers from the new study is quoted as saying, "If diet is the factor that lowers the sex ratio in vegetarians, it would appear that this effect overrides the effect of vegetarians' low smoking rate in pregnancy."

The second article also presented some historical beliefs about what affects the sex of a child. At the beginning of the 20th century, "it was firmly believed that an older father, or one married to a very strong, domineering woman, was likely to have boys." The fact that servicemen returning from World War II tended to father more boys than other men was seen as supporting this idea. Further support was found in studies of Burke's Peerage. "The habit for centuries was for the rich to take a bride considerably younger than themselves, whereas in poor families marriage was earlier, and the ages closer. As recorded in Burke's, the aristocracy had more sons."

"Putting Rankings to the Test: US News's College Evaluations Questioned"

by Jay Mathews, *The Washington Post*, 25 August 2000, C1.

The Post describes the annual *US News & World Report* college rankings issue as "a source of gastrointestinal distress for college presidents, alumni fundraisers and professional statisticians." The article cites a 1997 report prepared for *US News* by the National Opinion Research Center (NORC) in Chicago. Echoing some long-standing concerns of outside critics, the report states that "the principal weakness of the current approach is that the weights used to combine the various measures into an overall rating lack any defensible empirical or theoretical basis." The selection of factors to include in the formulas was also criticized. For example, there are no direct measures of quality of the curriculum or of student experience.

The Post received a copy of the NORC report from Nicholas Thompson of the *Washington Monthly*. Thompson's own commentary, entitled "Playing With Numbers: How U.S. News Mismeasures Higher Education and What We Can Do About It," can be read on-line at

<http://www.washingtonmonthly.com/features/2000/0009.thompson.html>

At this site you can find the NORC report and a response from *US News*. You will also find the text of a famous 1997 letter from former Stanford president Gerhard Casper, who wrote to *US News*: "I hope I have the standing to persuade you that much about these rankings -- particularly their specious formulas and spurious precision -- is utterly misleading." One of the primary concerns of the NORC report is the lack of research on the statistical properties of the *US News* estimators. According to NORC, *US News* has not even performed a simple correlation analysis on the components of its formulas.

Thompson himself laments the "beauty contest" mentality promoted by the rankings issue. He further worries about what he calls the "Heisenberg effect"; that is, colleges may actually institute changes in response to the *US News* rating categories. Admissions officials are certainly aware of the power of the rankings. They see application rates and yield go up when their school rises in the rankings; drops in the rankings have the opposite effect. Thompson cites findings by the National Bureau for Economic Research (NBER) that quantify the yield effect: to compensate for the decreased yield resulting from a one-place drop in the rankings, a school would need to increase its admissions rate by 0.4 the following year.

Readers of last year's rankings may recall Caltech's surprising leap to the number one spot among national universities. Thompson provides some details on how this came about. The pivotal change came in the spending-per-student measure. In prior years, it had entered the formula through ranks. But last year, absolute dollar amounts were used. To illustrate, Thompson cites data from 1997. Caltech was the runaway leader, spending \$74,000 per student. Yale was fourth with \$45,000 and Harvard seventh with \$43,000. In terms of ranks, Caltech's three-position advantage over Yale is the same as Yale's lead over Harvard. But in absolute numbers, Caltech spent 40% more than Harvard, whereas Harvard and Yale were nearly the same. The switch to absolute dollars gave Caltech a dramatic advantage last year. Thompson reports that this change caused serious internal disputes at *US News*.

"The Wild Poll Pendulum"

by William Safire, *The New York Times*, 28 September 2000, A27.

In this essay, Safire protests the unreasonable volatility of polling results, and the way that news stories are crafted to respond to them. He questions whether Al Gore's celebrated kiss with his wife Tipper really propelled him into the lead, only to be countered when George Bush kissed Oprah Winfrey on her talk show.

One source of difficulty, says Safire, is the pressure the media feel to actually have a story about which candidate is "winning" at any given point in time. Not content to report a large "undecided" number, polling organizations feel compelled to follow up with questions about who voters are "leaning towards." The flip side of this coin is that people do not like to appear uninformed, so they are easily pushed into taking some sort of stand. Safire wonders how much of the perceived volatility in the electorate is simply the result of changes in such "leanings" among voters who have legitimately not made up their minds.

Safire is, of course, famously concerned with precise use of language. Here he worries about what is left unsaid when polling results talk about the number of "respondents." He states that in 1984, a 65% response rate was typical, whereas today some of his friends in the polling community quietly admit that the rate is now as low as 35%. He suspects that the use of answering machines or caller ID is allowing potential respondents to screen their calls.

As I write this, the election is over, though we still do not know who the next president will be! In hindsight, maybe it is not surprising that the polls were not able to pin down a winner.

The debate about how to interpret the results in Florida is likely to continue for some time. Here is a web site that has catalogued a number of the fascinating statistical analyses that have been done to date:

<http://madison.hss.cmu.edu/>

"Failed Firestone Tires Seem Worse on Rear Driver's Side"

by Justin Hyde (Associated Press), *The Milwaukee Journal Sentinel*, 24 September 2000, A1.

The article reports on analysis done by Keith Baumgardner, a consultant who is studying data on Firestone tires in connection with lawsuits that have been brought against that tire-maker and Ford Motor Company. Baumgardner has seen 63 cases of tread separation on Ford Explorers. Of these, 45 involved rear tires -- 27 on the driver's side and 18 on the passenger's side. The other 18 either involved front tires or their position was not known.

Ford spokesman Ken Zino acknowledged that the company is aware of a slight trend toward rear tire failure and is working to identify the cause. Another Ford spokesman, Mike Vaught, reported that several theories were being considered:

The location of the Explorer's fuel tank on the left side might put more weight on that wheel.

The rotation of the Explorer's drive shaft might place more force on the left side than the right.

American roads might radiate more heat from the center of the pavement than from the edges, increasing the heat transferred to the driver's side tires.

Vaught stressed, however, that the investigation was still in its early stages, and that these theories did not contradict Ford's position that the tires themselves were to blame, not the design of the Explorer.

You can find additional information on the National Highway Traffic Safety Administration (NHTSA) web site:

<http://www.nhtsa.dot.gov/hot/Firestone/Update.html>

There you can find a database of complaints regarding Firestone ATX and Wilderness tires. As of an October 24 update, there were 3700 reports. The information is downloadable in Microsoft Excel format.

"Studies Find Race Disparities in Texas Traffic Stops"

by Jim Yardley, *The New York Times*, 7 October 2000, A12.

In recent years, the controversial use of racial profiling on New Jersey highways has received wide media attention. The present article describes studies conducted in Texas by the NAACP and by the *Dallas Morning News*. Both examined the rates at which African-Americans, Hispanics, and whites were ticketed and searched on Texas roadways. The *Morning News* looked at the 859,000 traffic tickets issued in Texas over the last year, while the NAACP focused on 65,000 drivers who were stopped during the month of March. Their findings were similar. According to the article, "Black and Hispanic motorists across Texas are more than twice as likely as non-Hispanic whites to be searched during traffic stops, while black drivers in certain rural areas of the state are also far more likely to be ticketed..."

Although the *Dallas Morning News* found that statewide, African-Americans and Hispanics received tickets "at rates that were proportional to their driving-age populations," they also concluded that in many rural counties, blacks were ticketed nearly twice as often as non-Hispanic whites. In response, the Texas Department of Public Safety said the study was flawed because, as the article states:

It compared the race and number of ticketed drivers with the local population where the stop occurred, but didn't consider that the drivers might be from elsewhere.

The state's own analysis reportedly shows that non-Hispanic white drivers are stopped more often "than their estimated state wide population," but that blacks and Hispanics are twice as likely to be searched.

[JSE Homepage](#) | [Subscription Information](#) | [Current Issue](#) | [JSE Archive \(1993-1998\)](#) | [Data Archive](#) | [Index](#) | [Search JSE](#) | [JSE Information Service](#) | [Editorial Board](#) | [Information for Authors](#) | [Contact JSE](#) | [ASA Publications](#)