SUPERIOR COURT OF DECATUR COUNTY
STATE OF GEORGIA

JAMES BRYAN WALDEN and	*	
LINDSAY WALDEN, Individually and	*	
on Behalf of the Estate of Their Deceased Son,	*	
REMINGTON COLE WALDEN,	*	
	*	
Plaintiffs,	*	
	*	CIVIL ACTION
v.	*	
	*	FILE NO. 12-CV-472
CHRYSLER GROUP, L.L.C., and	*	
BRYAN L. HARRELL,	*	
	*	
Defendants.	*	

AFFIDAVIT OF NORMA F. HUBELE, PH.D.

STATE OF ARIZONA
COUNTY OF MARICOPA

Personally appeared before the undersigned attesting officer, duly authorized to administer oaths, Norma F. Hubele, Ph.D., who testified as follows:

1.

I am over eighteen years old and am of sound mind. Under power of oath, I swear and affirm that the following is true and correct.

2.

The purpose of this affidavit is to summarize and illustrate my opinions about the underlying statistically flawed analysis performed by Dr. Paul M. Taylor (hereafter Taylor) and Dr. Laurentius Marais (hereafter Marais) in *Walden v. Chrysler*. The focus is on the methodologies used and why these methodologies do not constitute a sound, statistical basis for their opinions. Elsewhere in my deposition and exhibits of June 26, 2014, I have put forth my opinions about the appropriate comparison vehicles, rates and conclusions concerning the fatal rear impact crashes with fire as the most harmful event, as happened in the Walden vehicle.

3.

My opinions are based on my 25+ years of experience as a full professor in statistics, authorship of numerous articles and a best-selling statistical textbook, plus 25+ years of extensive work with the databases used by Taylor and Marais.

4.

The Fatality Analysis Reporting System (FARS) "is a nationwide census providing NHTSA, Congress and the American public yearly data regarding fatal injuries suffered in motor vehicle traffic crashes."[1] It therefore follows that all counts obtained from FARS are census counts. Census counts all have the same high reliability with no statistical errors. Taylor uses this census FARS database to construct his rates for the Jeep Grand Cherokee and his selected comparison vehicles of model years 1999-2004.

5.

Vehicle registration counts, called registered-vehicle-years, are also census counts.[2] Again, it therefore follows that all registered-vehicle-years among different comparison vehicles have the same high reliability and no statistical errors.

[1] Copied from the website http://www.nhtsa.gov/FARS accessed on December 10, 2014.

[2] Registered-vehicle-years is computed by summing the number of vehicles registered each year over a period of several years. For example, if 400,000 Ford Explorers were registered in 2009 and 530,000 were registered in 2010, then summed together we would report that there were 930,000 registered-vehicle-years for the Explorer for the period 2009-2010. One vehicle on the road during both 2009 and 2010 would be counted twice. This unit of "registered-vehicle-years" is a typical unit used as a proxy for "exposure" and is useful to normalize FARS counts of large and small volume vehicles.

6.

Taylor uses FARS counts as the numerator and registered-vehicle-years as the denominator to form rates. Since FARS counts and registered-vehicle-years are both census counts, then Taylor's rates are census rates with the same high reliability and no statistical errors. For ease of discussion, rates formed by the ratio of FARS counts and registered-vehicle-years will be called 'census FARS-registration rates' in this report.

7.

Since the census FARS-registration rates are not derived from survey samples, there is no need to perform statistical inference techniques.[3] Statistical inference techniques are methods used to *infer* information from a sample to the broader, *un-sampled population*. Since there is *no un-sampled population* in FARS, there is no need for statistical inference techniques.

8.

On the one hand, Taylor admits that FARS are census counts with no statistical errors, but then, on the other hand, he states that the "play of chance" makes it necessary to create *confidence intervals*, to accompany his rates (*Deposition of Paul Taylor* in *Walden v Chrysler*, October 2, 2014, p. 144, lines 7-14).

As will be discussed below, Taylor *assumes* a probability model, treats the fatalities in FARS as *sampled values* and constructs confidence intervals, i.e., a *range of rates*, for his random models. This approach results in misleading and deceptive conclusions.

---

[3] Dr. Mary Natrella, a well-respected statistician with the National Bureau of Standards in the early 1960's, wrote "*If we were willing or able to examine an entire population, our task would be merely that of describing that population, using whatever numbers, figures, or charts we care to use.*" *Experimental Statistics*, reprint of the Experimental Statistics Portion of the AMC Handbook, by Mary Natrella, p. 1-3, National Bureau of Standards Handbook 91, 1963, emphasis added. See Appendix A for more information about *Experimental Statistics* and the author.

9.

Confidence intervals are classical statistical inference techniques used in associated with sampled data. Based on the assumption that there are un-sampled data, these intervals are useful tools for bracketing single-value estimates and providing reliability information about the un-sampled data. A confidence interval gives a range of values within which the true unknown value is assumed to lie. Large amounts of un-sampled data or large variety in the un-sampled data will lead to a wide confidence interval and signal an unreliable estimate of the true unknown value. A narrow or short confidence interval indicates small amounts of un-sampled data, little variety in the data and a very reliable estimate.[4] [5]

10.

Taylor treats the census FARS-registration rates as sample estimates and constructs confidence intervals on his rates. While census FARS-registration rates are all equally highly reliable, Taylor's confidence intervals are extremely different in size and therefore give the misleading conclusion that the rates have different reliability.

For example, Figure A gives the length of Taylor's confidence intervals on his census FARS-registration rates for fatal rear impact vehicles with fire as the most harmful event, as experienced in the Walden vehicle. The length of his intervals measure from 0.24 for the Explorer to 4.4 for the Isuzu Trooper. In other words, the Isuzu Trooper confidence interval is

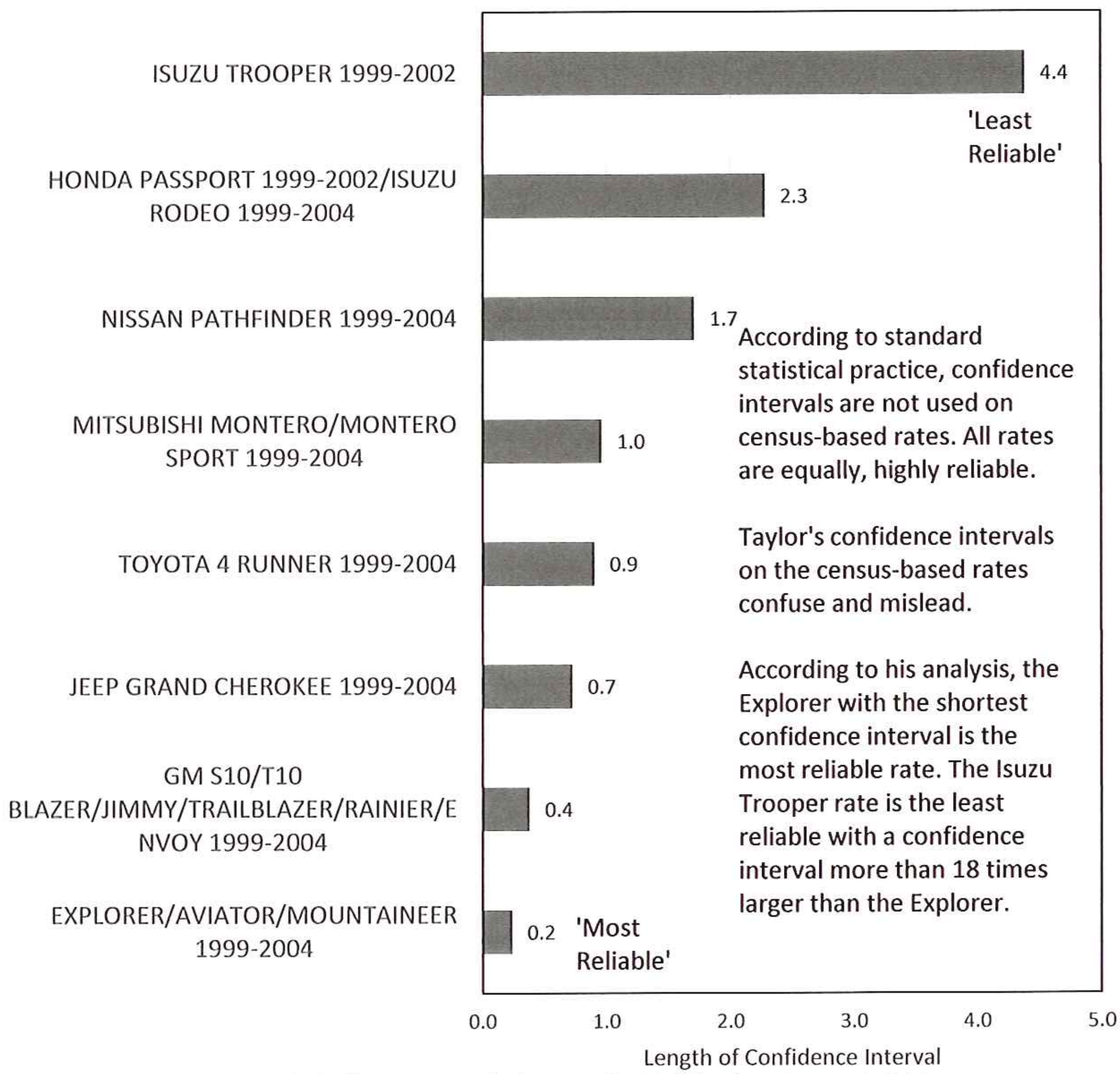
---

[4] Montgomery, D., Runger, G. and Hubele, N. F. (2011). *Engineering Statistics, 5th Edition*, Wiley & Sons, Chapters 4-8. Probabilities, such as 95%, are used to signal the *confidence* in the range of value for capturing the true value.

[5] As an example of an unreliable estimate, consider a voter's opinion sample which indicates that 56% of the voters favor a particular candidate, but the confidence interval of that percentage is [48%, 64%]. Such a wide confidence interval indicates that the candidate could narrowly lose the election with 48% of the vote or have a stunning win with 64% of the votes. The wide confidence interval signals an unreliable estimate, leading to an unreliable prediction. The confidence interval is so big as to be meaningless.

more than 18 times larger than the Explorer's. This large disparity in length leads to the misleading statistical conclusions that the Isuzu Trooper rate is less reliable than the Explorer rate.

Figure A. Length of Confidence Intervals on Taylor's Census FARS-Registration Rates for Fatal Rear Impact Vehicles with Most Harmful Event = Fire



Source: Taylor Excel File "Summary File for CI Calc QCd" in the electronic folder "Exponent for Walden." Calendar Years 1998-2012

11.

Taylor proceeds to use his misleading confidence intervals to draw conclusions about statistically significant differences among the vehicles that he compares. Since all the ranges of values in the confidence intervals overlap, again drawing on concepts in statistical inference, Taylor concludes that the Jeep Grand Cherokee rate is not statistically significantly different than any of the comparison vehicles (*Taylor Deposition*, p. 245, lines 10-25; p. 246, lines 1-14). Such statistical conclusions are based on a *mistreatment* of the census FARS-registration rates as sample values. As paraphrased from footnote 3, if we have all the data, our task is merely to compile charts and draw our own conclusions, independent of any statistical inference methods.

12.

Taylor attempts to shore up his mis-use of confidence intervals by providing a handful of publications that describe research endeavors and probability models developed for research with transportation data.[6]

When examined closely, however, none of the research papers use Taylor's data, rates or methodology. Specifically, none of the researchers use the census FARS-registration rates and none of them use the probability models of Taylor. His use of these references indicates that he is not aware of the important difference between his methodology and those cited. Furthermore, during a federal defect investigation presentation, Chrysler explicitly cites one research report to

[6] "The Long-Term Effects of ABS in Passenger Cars and LTVs" DOT HS 811 182; "Relationships between Vehicle Size and Fatality Risk in Model Years 1985-93 Passenger Cars and Light Trucks" DOT HS 808 570; "A Preliminary Evaluation of Two Braking Improvements for Passenger Cars Dual Master Cylinders and Front Disc Brakes" DOT HS 806 359; "An Evaluation of Side Marker Lamps for Cars, Trucks and Buses" DOT HS 806 430; "An Analysis of Traffic Deaths by Vehicle Type and Model" referenced by Taylor as "Confidence Interval SUV-report LBNL Report for DOE."

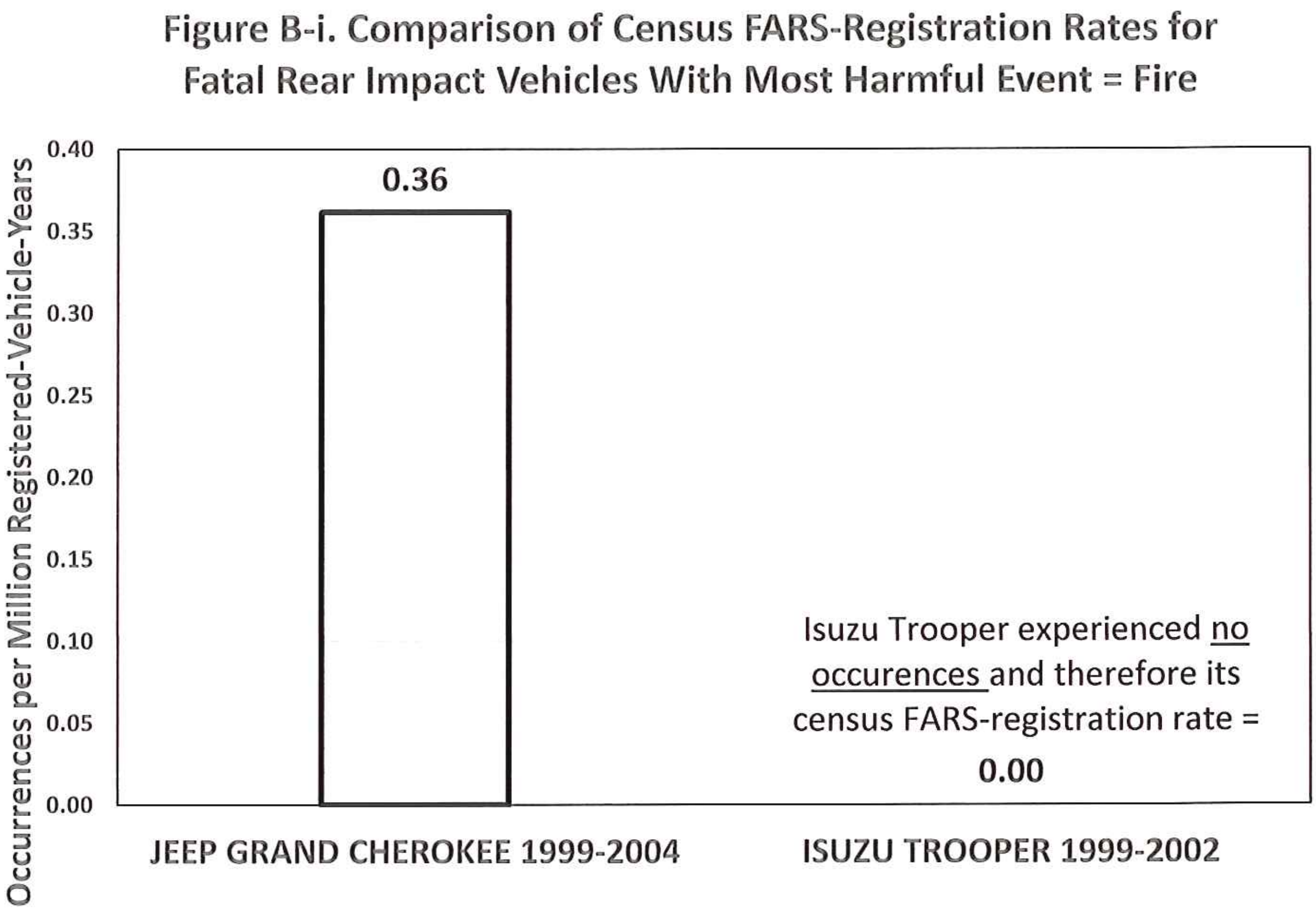
justify Taylor's probability models and confidence intervals. A close read of that report specifically rejects the methodology chosen by Taylor and Chrysler.[7]

13.

When asked at his deposition for his reasons for using confidence intervals, Taylor is vague and unscientific. He never says that he is making an important underlying assumption, i.e., he has chosen a specific probability model to construct his confidence intervals (*Taylor Deposition*, p. 142, lines 16-25, p. 143, line 1). Taylor cites Marais when he mentions "play of chance" and changing "individual circumstances" as reasons for using confidence intervals. However, as the following figures will demonstrate, his ambiguous reasons for using confidence intervals on census FARS-registration rates lead to an inappropriate adoption of a theoretical model, resulting in unsound and misleading results.

As shown in Figure B-i, according to Taylor's data, the Jeep Grand Cherokee had 5 occurrences of fatal rear impacts with fire as the most harmful event yielding a rate of 0.36 per million-registered-vehicle-years. The Isuzu Trooper had no occurrences and therefore had a census FARS-registration rate of 0.

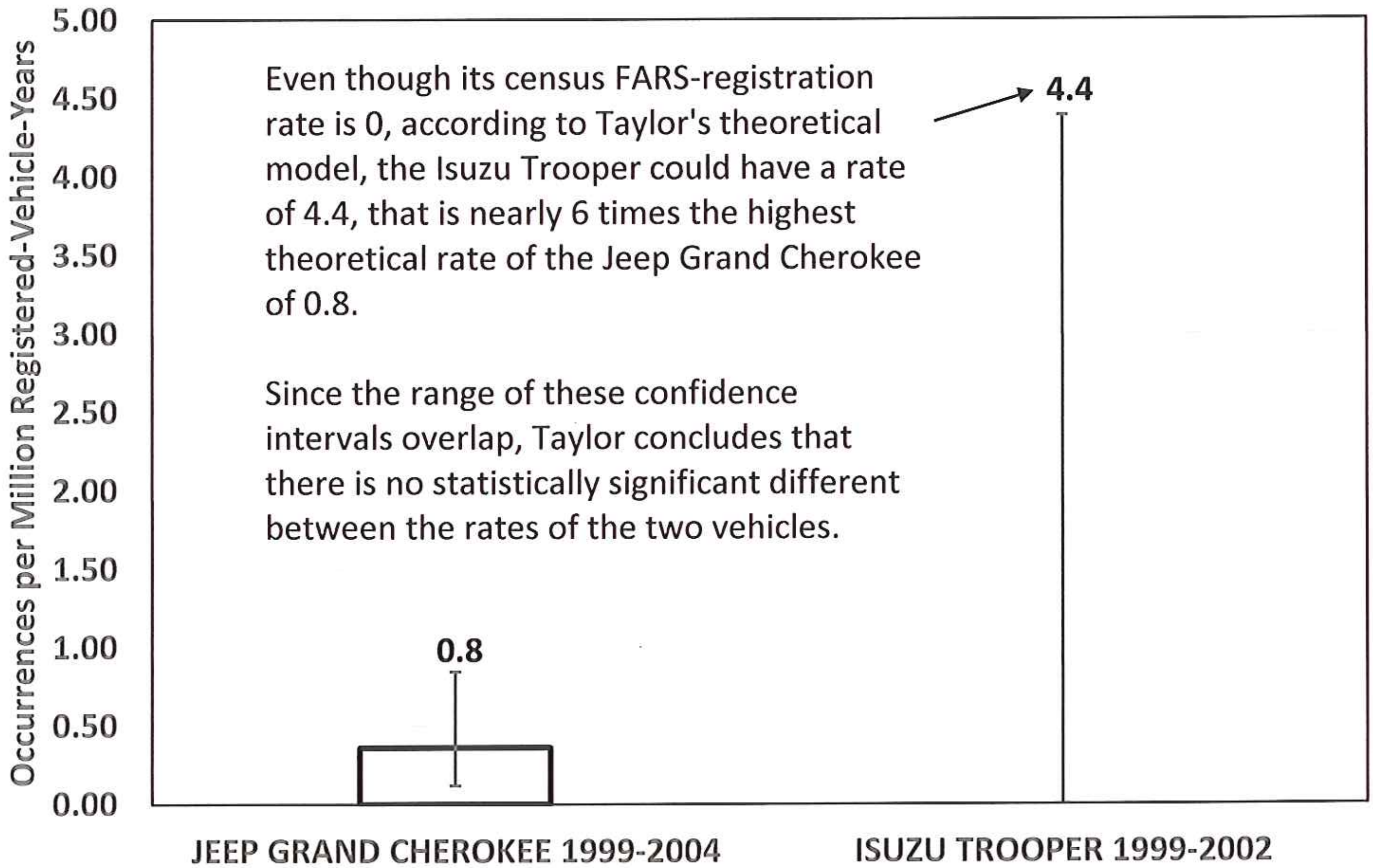
[7] In order to construct their confidence intervals, Taylor and Chrysler assumed a probability model to theoretically describe the occurrences of crashes. In Chrysler's August 29, 2012 presentation entitled "EA12-005 Relevance of Confidence Intervals to the Analysis of FARS Data" the research paper DOT HS 811 572 p. 4 is cited as the justification for their choice of the Poisson distribution probability model: "The negative binomial and Poisson distributions are commonly used in modeling crashes over time or miles." (See Appendix B at the end of this report.) However, deep in the DOT HS 811 572 report we find that the author specifically chose the negative binomial model instead of the Poisson: "The dispersion parameter is significantly different from (and larger than) zero, confirming our choice to model crashes via a negative binomial distribution, <u>instead of Poisson</u>." (emphasis added, p. 20) Notably, the DOT author cited by Chrysler explicitly rejects the model chosen by Taylor and Chrysler.



Source: Taylor Excel File "Summary File for CI Calc QCd" in the electronic folder "Exponent for Walden." Calendar years 1998-2012.

Figure B-ii shows Taylor's confidence intervals superimposed on the rates. The upper value of the Jeep Grand Cherokee confidence interval is 0.8, whereas the Isuzu Trooper's upper value is 4.4. That is, according to Taylor's model, the Trooper could have nearly 6 times the rate of occurrence as the Jeep Grand Cherokee, even though none have occurred in the Trooper, to date. This result renders the confidence intervals meaningless and misleading.

Figure B-ii. Taylor's 95% Confidence Intervals on
Census FARS-Registration Fatal Rear Impact Vehicles Rates with
Most Harmful Event = Fire

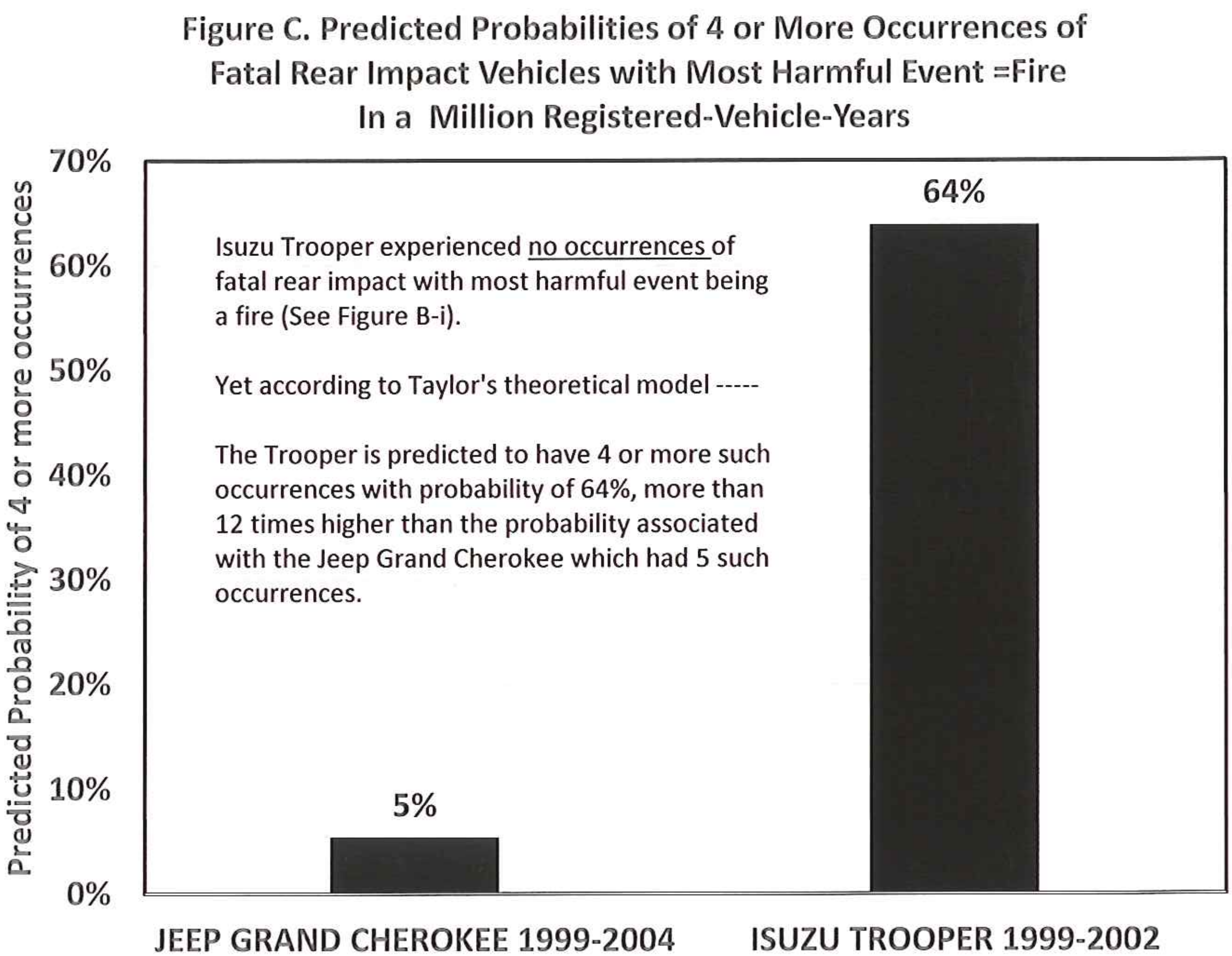


Source: Taylor Excel File "Summary File for CI Calc QCd" in the electronic folder "Exponent for Walden." Calendar years 1998-2012.

14.

In Figure C, we show that Taylor's probability models lead to more misleading conclusions. According to Taylor's upper limit of his confidence interval (with a value of 0.8) and his probability model, the Grand Cherokee has a 5% chance of 4 or more occurrences of a fatal rear impact crash with fire as the most harmful event happening in a million-registered-vehicle-years. In contrast, using his upper limit for the Trooper (4.4), his probability model predicts that the Trooper has a 63% chance of 4 or more occurrences. In other words, when 63%

is compared to 5%, this indicates that according to Taylor's theoretical models, a vehicle which has had no occurrences is predicted to have more than 12 times higher probability of 4 or more occurrence than the Jeep Grand Cherokee. This misleading result is a consequence of assuming an inappropriate theoretical model and treating the census FARS-registration rates as sample statistics.



Source: Using the upper bound of the theoretical Poisson model applied to the FARS 1998-2012 database.

15.

Taylor uses these flawed analyses to opine that there are no statistically significant differences among the various comparison vehicles and crashes. His statistical analysis is not based on standard statistical practice and leads to misleading and deceptive results. Such analysis does not constitute a sound basis for his opinion in this matter.

16.

Chrysler's second expert to render statistical opinions, Marais, clearly states in his deposition that he did not do any statistical work specifically for this case (*Deposition of Laurentius Marais* in *Walden v Chrysler*, August 27, 2014, p. 58, 9-14). This lack of focus is evident in his very broad graphical displays contained in Exhibits 15 and 16 of his deposition.

17.

In contrast to the census FARS database, Marais uses two federal sample databases, the National Accident Sampling System General Estimation System (NASS GES or simply GES) and the National Accident Sampling System Crashworthiness Data System (NASS CDS or simply CDS).[8] Of the approximately 5.5 million crashes annually, the GES samples only about 50,000, i.e., less than 1.0% annually. The CDS system samples only about 5,000, i.e., less than 0.1% annually.[9]

18.

With such small percentages of national crashes, these databases are not intended to be nationally representative of the crash experience of individual vehicles. In essence, the sample

---

[8] The GES is a scientific survey sample of police accident reports. The NASS CDS is a different scientific survey sample containing detailed information, over 200 data elements, about the crash compiled by experienced investigators who study the crash scene, the vehicle, occupant medical reports, as well as the police accident report. See Appendix C for a brief description of the two databases.

[9] In 2010 there were nearly 5.5 million police-reported crashes on the nation's roadways, as reported in NHTSA's *Traffic Safety Facts*, 2010, http://www-nrd.nhtsa.dot.gov/Pubs/811659.pdf

sizes (1.0% and 0.1% of all crashes) are just too small to generate reliable estimates for individual vehicles or small groups of vehicles in specific types of crashes. Marais violates this standard statistical practice when he used the GES and CDS as a basis for his opinions. [10]

Table 1 lists the small group of vehicles contained in the "Subject Vehicles" used by Marais and the specific types of crashes selected from the GES and CDS databases. For example, his search of the GES and CDS databases for crashes of subject vehicles that experienced post-collision fire in rear impacts resulted in finding 3 and 2 vehicles. The 'largest' sample size for his estimates on rear impact crashes consists of 3 vehicles. Such extremely small sample sizes are a result of 'drilling' too far down into the database. These extremely small sample sizes result in very unreliable estimates that do not provide a sound basis for his opinions.

Criteria Marais Used to Select Crashes from the Databases for "Subject Vehicles" (Jeep Grand Cherokee 1993-2004, Jeep Liberty 2002-2007, Jeep Cherokee 1993-2001)	**Sample Size: Number of Vehicles Selected from Database**	
	GES	**CDS**
Occurrence of Post-Collision Fire in Rear Impacts	3	2
Occurrence of Most Harmful Event = Fire in Rear Impacts	1	Not Coded
Occurrence of Post Collision Fire with Origin in Fuel Tank Area	Not Coded	3
Occurrence of Post-Collision Fire with Origin in Fuel Tank Area in Rear Impacts	Not Coded	2
Occurrence of Post-Collision Fuel System Leakage	Not Coded	6
Occurrence of Post-Collision Fuel System Leakage in Rear Impacts	Not Coded	3

Table 1. Extremely Sample Sizes Used by Marais to Construct GES and CDS Graphs Listed in his of Exhibits 7, 15 and 16. "Not Coded" indicates that the database does not contain this information and therefore was not used.

[10] See any sampling theory textbook for discussion of the importance of large samples in computing reliable statistics, e.g., S. Lohr, *Sampling: Design and Analysis*, Pacific Grove: Duxbury, 1999.

19.

Marais's comparison vehicles contrast sharply with those put forth by Taylor.[11] On the one hand, Taylor argues to restrict the focus of his analysis to comparison vehicles of model years 1999-2004, and, on the other hand, Marais uses vehicles spanning model years 1984-2005. On the one hand, Taylor restricts his comparison vehicles to a select handful of SUVs and, on the other hand, Marais uses broad groups of SUVs and cars. Also, Marais does not restrict his SUVs by size; he includes large and small SUVs among his comparison vehicles. Finally, Marais chooses to use cars as a comparison vehicle group. This last group is so broad as to include two-door coupes, hatchbacks and sedans. From a statistical point of view, any conclusions based on statistical comparisons of these broad groupings of vehicles would be too general to draw conclusions about the specific performance in rear impacts of the 1999 Jeep Grand Cherokee at issue in this case.

20.

As demonstrated in Table 1, Marais uses extremely small samples to arrive at his estimates. Also, as described in the Taylor discussion, large confidence intervals signal the unreliability of an estimate. An examination of the graphs in Marais's Exhibits 7, 15 and 16 demonstrate that in almost every instance the confidence interval for the "Subject Vehicles" is nearly twice the length of the bar graph. There are two primary reasons for these large confidence intervals: the sample sizes are exceedingly small and the standard errors are exceeding large.

[11] The definition of the individual groups can be found in the footnote of Figures D of this affidavit, reproduced from a graph found in Marais's Exhibit 16 in *Walden v Chrysler*.

For every estimated rate computed by Marais from the GES and CDS sample databases, he also computed a standard error of that rate. The standard error is, in turn, used to construct the confidence intervals on his rates. To gauge the reliability of an estimate it is useful to compare the standard error to the estimate. In fact, the ratio of the standard error to the estimate is called the coefficient of variation. Since the standard error also incorporates the sample size, the coefficient of variation is a useful metric for assess the adequacy of the sample size for providing reliable estimates.

If the coefficient of variation exceeds about 10%, then this signals that the standard error is too large and the sample size inadequate for producing reliable statistics for rates.[12] All of Marais's estimates for all his GES and CDS rates for his "Subject Vehicles" substantially exceed the 10% recommended cutoff. In fact, his coefficients of variation are in excess of 50%.

Visually, these large coefficients of variation resulting from extremely small sample sizes and large standard errors translate into the large confidence intervals found in Exhibits 7, 15 and 16 for the "Subject Vehicles." The heights of the bars in Marais's graphs are his estimated rates. His confidence intervals are superimposed on the bars.[13]

Figure D is a reproduction of Marais's graph for CDS-based rates of post-collision fuel system leakage in rear impacts. The confidence interval for "Subject Vehicles" extends from 0 to 220. Such a large interval is based on only 3 vehicles, as shown in Table 1. The coefficient of variation is 67%. Together, the exceedingly small sample size, the large coefficient of variation and the resultant large confidence interval of the "Subject Vehicles" indicate a very unreliable

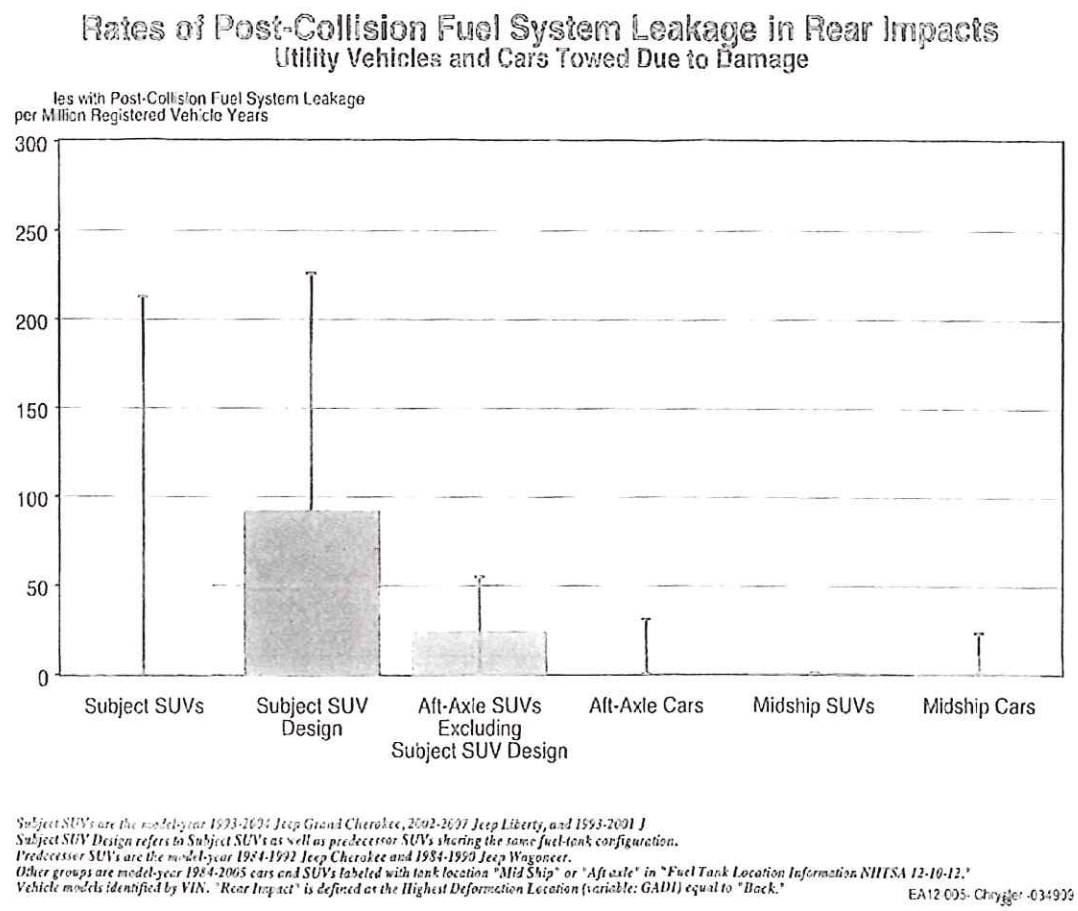
---

[12] See W. Cochran, *Sampling Techniques*, 3rd Edition, New York: Wiley, p. 53-55.

[13] The estimates, standard errors, coefficients of variation and confidence intervals derived from the GES and CDS databases are based on a mathematical algorithm. Since the GES and CDS are complex samples, as opposed to simple random samples, a mathematical algorithm performs the necessary weighting and compiling of information to arrive at these values.

estimate. Marais uses this large confidence interval and its inevitable overlap with all the other confidence intervals of the other groups as the basis for his conclusion that there is no statistically significant difference between the rates.

Figure D. Copy of One of Marais's CDS Graphs from Exhibit 16 With Very Wide Confidence Intervals Indicating Very Unreliable Estimates



Furthermore, with a lower value of 0 for the confidence interval, Marais is saying that the true rate could be 0. However, we already know that the Walden vehicle had a post-collision fuel system leak in a rear impact, therefore, the rate is truly not 0. This only adds more indications that the confidence intervals generated by Marais are unreliable and meaningless.

21.

These unreliable rates, large coefficients of variation, large confidence intervals, extremely small samples and broad groupings of vehicles do not provide a sound statistical basis for any opinions by Marais about the specific performance of a 1999 Jeep Grand Cherokee at issue in this case.

22.

My opinions were formulated based on the depositions and exhibits of Taylor and Marais in this matter and my experience as a statistician and scholar. These criticisms extend beyond the individual examples cited and would be applicable to other charts of a similar nature concerning FARS, GES and CDS data.

This 19 day of December, 2014.

Norma F. Hubele

NORMA F. HUBELE, PH.D.

I, Jameil Palmer, a notary public for the county and state, certify that Norma F. Hubele, Ph.D., whose name is signed to the foregoing instrument and who is known to me, acknowledged before me on this day that, being informed of the contents of the instruments, he executed the same voluntarily.

Sworn to and subscribed before me this 19th day of December, 2014.

JAMEIL PALMER
Notary Public - Arizona
Maricopa County
My Comm. Expires Jan 26, 2015

Appendix A. Short Biographical Information on Mary Gibbons Natrella

Experimental Statistics

NBS Handbook 91: Experimental Statistics [1] was first published in 1963. The material for the book was originally commissioned and printed in a limited edition by the U.S. Army as a series of five Army Ordnance Pamphlets OSRDDP 20-110-114. The publication was prepared in the Statistical Engineering Laboratory (SEL) of NBS under a contract with the former Office of Ordnance Research. Although originally intended for the needs of the Army, it proved to be equally useful to research and development groups, both within and outside the government.

The Handbook is organized in five sections; namely,

1. Basic Statistical Concepts and Analysis and Interpretation of Measurement Data
2. Standard Techniques for Analysis and Interpretation of Enumerative and Classificatory Data
3. The Planning and Analysis of Comparative Experiments
4. Special Topics
5. Tables

It is a monumental work which brings together under one cover the combined experience and expertise of the Statistical Engineering Laboratory. Mary Gibbons Natrella was principal author with overall responsibility for the entire publication, but there are contributions on polynomial and multivariable relationships, sensitivity testing, use of transformations, and expression of uncertainties by other members of the SEL. Chapters 1, 20, and 23 were written by Churchill Eisenhart, founding father of the SEL; Chapter 6 was written by Joseph Cameron, who later became Chief of SEL, and Chapter 10 is based on material prepared by Mary Epling. Some original tables were prepared by Paul Somerville; Norman Severo assisted with Section 2, and Shirley Young Lehman helped with the collection and analysis of examples.

Mary Natrella had a special gift for elucidating difficult statistical concepts, and these expositions are the strength of the book. The workbook style of the volume probably accounts for its popularity and acceptance by statisticians and non-statisticians alike. It is replete with examples; the page for each example is divided, with the statement of the problem and recommended solution on the left-hand side and detailed step-by-step calculations on the right-hand side. Mary

Fig. 1. Mary Gibbons Natrella.

Natrella also believed in attention to detail. The Foreword states that "some procedures in the Handbook have been explained and illustrated in detail twice: one for the case where the important question is whether the performance of a new material, product, or process exceeds an established standard; and again for the case where the important question is whether its performance is not up to the specified standard."

The Handbook was an immediate success at NBS, in the Army, and throughout the Department of Defense. It eventually received wide acclaim in other government agencies, industry, and universities. Churchill Eisenhart was fond of quoting a statistician who said that "the best thing about the Handbook is that it is correct."

The Handbook is recognized for its deep and long-lasting impact on the application of statistics to the planning and analysis of scientific experiments. It was reprinted in 1983 for commercial sale by Wiley Interscience as part of its Selected Government Publications series. In 1985, the American Society for Metals published a condensation of four chapters on planning and analysis of comparative experiments as part of the Statistics Section of Volume 8 of the 9th edition of the ASM Handbook. It has been NIST's second-best selling publication, after the *Handbook of Mathematical Functions,* which is covered elsewhere in this volume. The material is still current after more than thirty years, and this year alone it received close to forty journal citations as measured by the Science Citation Index.

NIST still receives requests for this book, and its contents are the basis for training courses taught by the Statistical Engineering Division (SED) and companies such as SEMATECH that are involved in technology development. It has proved such an inspiration that a few years ago Patrick Spagon of the Statistical Methods Group of SEMATECH approached SED with a proposal for updating and recreating the book with examples directed towards the semiconductor industry. That proposal has evolved into a publication for the World Wide Web [2] that is currently under development by a team that includes: James Filliben, William Guthrie, Alan Heckert, and Carroll Croarkin of SED; Paul Tobias, head of the Statistical Methods Group, and Chelli Zey of SEMATECH; Barry Hembree of AMD; and Ledi Trutna, a private consultant.

Mary Natrella joined the Statistical Engineering Laboratory of NBS as a mathematical statistician in April 1950 after Churchill Eisenhart noticed her work as a sampling inspection expert in the U.S. Navy's Bureau of Ships. At the Navy, she worked on a team that developed the now famous MIL-STD-105A, Sampling Procedures and Tables for Inspection by Attributes (1950), which was to become a government wide standard. Her "Report of Proceedings of the Subcommittee for Preparation of MIL-STD-105A," issued as a companion document to the standard, was distributed to government groups developing sampling inspection plans, as well as to professional and technical organizations, and was still in demand in the 1970s.

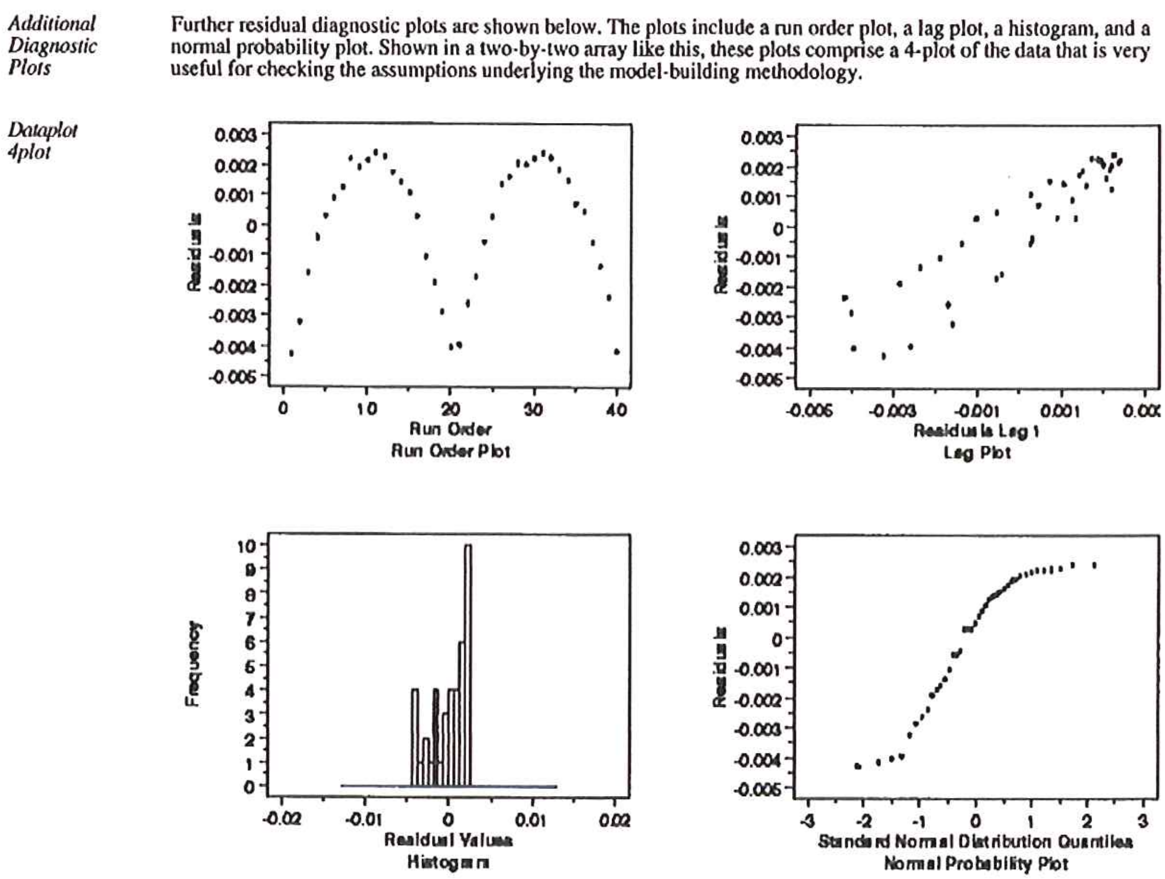


Fig. 2. Partial page from *NIST/SEMATECH Engineering Statistics Handbook* showing graphical techniques for deciding if the relationship between two variables is linear.

The SEL was a pioneering group dedicated to application in the physical sciences of the statistical methods developed by R. A. Fisher, J. Neyman, and E. S. Pearson in England and Walter Shewhart in the United States. Mary Natrella served as a consultant on statistical planning and analysis of experiments to NIST scientists, and her training course on "Statistics of Measurement" was heavily attended for many years. Her most notable publication is *Handbook 91*, which was 7 years in preparation and required all her talents as teacher and consultant. During this time, Mary also contributed several sections to *NBS Special Publication 300* [3] and worked with Carroll Brickenkamp and Steve Hasko of the NIST Office of Weights and Measures on *NBS Handbook 133: Checking the Net Contents of Packaged Goods* [4], which has been adopted by the National Conference on Weights and Measures. She performed her last service for the Statistical Engineering Division by serving as Acting Division Chief for the 2 years prior to her retirement in 1986.

The 1980's brought the culmination of Mary's career, with recognition on many fronts. She was elected a Fellow of the American Statistical Association (ASA) in 1981. In 1982, she received the Department of Commerce's Superior Federal Service Award. She was a long-time member of American Society for Testing Materials (ASTM) Committee E-11 on Quality and Statistics. In 1984, for her work as chairman of Subcommittee E11.03 on Statistical Analysis and Control Techniques, she was awarded the Society's Award of Merit, which carries the designation of Fellow.

In the year 2000 Mary Natrella has been further honored by the establishment of an endowed scholarship fund which will provide $1000 scholarships each year for two students to attend the Quality and Productivity Research Conference (QPRC). The scholarships were established by the Quality and Productivity section of the American Statistical Association, under the chairmanship of Veronica Czitrom of Lucent Technologies, with funds donated by Mary's husband, Joseph V. Natrella, and the QPRC Steering Committee. The purpose is to honor Mary's 36 years as author, teacher, and consulting statistician and her many contributions to the statistical community.

Prepared by M. Carroll Croarkin.

Bibliography

[1] Mary Gibbons Natrella, *Experimental Statistics*, NBS Handbook 91, National Bureau of Standards, Washington, DC (1963); reprinted 1966.

[2] Paul Tobias and Carroll Croarkin (eds.), *NIST/SEMATECH Engineering Statistics Handbook*, (http://www.itl.nist.gov/div898/handbook/index.html), National Institute of Standards and Technology (1999).

[3] Harry H. Ku (ed.), *Precision Measurement and Calibration: Statistical Concepts and Procedures*, NBS Special Publication 300, Vol. 1, National Bureau of Standards, Washington, DC (1969).

[4] C. S. Brickenkamp, S. Hasko, and M. G. Natrella, *Checking the Net Contents of Packaged Goods*, NBS Handbook 133, Third Edition, National Bureau of Standards, Gaithersburg, MD (1988).

Appendix B: Excerpt from Chrysler Presentation Material

EA12-005
Relevance of Confidence Intervals to the Analysis of FARS Data

8/29/2012

EA12-005 (June 18, 2013) - Chrysler - 080

CGLLC042430

Method of Determining Confidence Intervals

- The confidence intervals shown in Examples #1 and #2 are based on the Poisson probability distribution, which reasonably describes the probabilistic nature of accident frequencies
- The Poisson distribution is commonly used for this purpose, as is the negative binomial distribution:
  - "The negative binomial and Poisson distributions are commonly used in modeling crashes over time or miles." (DOT Report HS 811 572, June 2012, p. 4)
- The binomial distribution used in previous Chrysler submissions closely approximates the Poisson distribution in the case of infrequent events

Jeep CHRYSLER DODGE RAM FIAT

EA12-005 (June 18, 2013) - Chrysler - 087

CGLLC042437

Appendix C. National Automotive Sampling System

General Estimation System Overview (GES)

Data for GES come from a nationally representative sample of police reported motor vehicle crashes of all types, from minor to fatal. The system began operation in 1988, and was created to identify traffic safety problem areas, provide a basis for regulatory and consumer initiatives, and form the basis for cost and benefit analyses of traffic safety initiatives. The information is used to estimate how many motor vehicle crashes of different kinds take place, and what happens when they occur. Although various sources suggest that about half the motor vehicle crashes in the country are not reported to the police, the majority of these unreported crashes involve only minor property damage and no significant personal injury. By restricting attention to police-reported crashes, GES concentrates on those crashes of greatest concern to the highway safety community and the general public.

GES data are used in traffic safety analyses by NHTSA as well as other DOT agencies. GES data are also used to answer motor vehicle safety questions from Congress, lawyers, doctors, students, researchers, and the general public.

GES Sampling

In order for a crash to be eligible for the GES sample a police accident report (PAR) must be completed, it must involve at least one motor vehicle traveling on a traffic way, and the result must be property damage, injury, or death.

These accident reports are chosen from 60 areas that reflect the geography, roadway mileage, population, and traffic density of the U.S. GES data collectors make weekly visits to approximately 400 police jurisdictions in the 60 areas across the United States, where they randomly sample about 50,000 PARs each year. The collectors obtain copies of the PARs and send them to a central contractor for coding. No other data are collected beyond the selected PARs.

Trained data entry personnel interpret and code data directly from the PARs into an electronic data file. Approximately 90 data elements are coded into a common format. Some element modification takes place every other year in order to meet the changing needs of the traffic safety community. To protect individual privacy, no personal information, such as names, addresses, or specific crash locations, is coded. During coding, the data are checked electronically for validity and consistency. After the data file is created, further quality checks are performed on the data through computer processing and by the data coding supervisors.

An annual publication, Traffic Safety Facts, is produced with GES data for nonfatal crashes, combined with information on fatal crashes from the Fatal Analysis Reporting System.

(Content copied from NHTSA website: http://www.nhtsa.gov/Data/National+Automotive+Sampling+System+(NASS)/NASS+General+Estimates+System on December 10, 2014)

Crashworthiness Data System Overview (CDS)

National Automotive Sampling System Crashworthiness Data System (NASS CDS or CDS) has detailed data on a representative, random sample of thousands of minor, serious, and fatal crashes. Field research teams located at Primary Sampling Units (PSU's) across the country study about 5,000 crashes a year involving passenger cars, light trucks, vans, and utility vehicles. Trained crash investigators obtain data from crash sites, studying evidence such as skid marks, fluid spills, broken glass, and bent guard rails. They locate the vehicles involved, photograph them, measure the crash damage, and identify interior locations that were struck by the occupants. These researchers follow up on their on-site investigations by interviewing crash victims and reviewing medical records to determine the nature and severity of injuries.

Interviews with people in the crash are conducted with discretion and confidentiality. The research teams are interested only in information that will help them understand the nature and consequences of the crashes. Personal information about individuals - names, addresses, license and registration numbers, and even specific crash locations - are not included in any public NASS files.

The data collected by the PSU's are quality controlled by one of 2 NASS Zone Centers. Each Zone Center, staffed by the most experienced crash researchers, is responsible for half of the PSU field offices. Zone Centers have the responsibility for coordinating and supervising the activities of the field offices, keeping field offices informed regarding changes in functional and administrative procedures, sharing ideas and concepts throughout the system regarding new techniques, procedures, and components found on vehicles and updating field offices regarding changes in system hardware and software.

NASS case review is conducted at the Zone Center and may result in case data being sent back and forth between the Zone Center and the PSU several times until the case passes quality control standards built into the NASS data collection cycle. Once data is approved for inclusion into the NASS database, it will again be subjected to quality assurance checks before becoming publicly released as part of annual NASS data files.

The data collected by the CDS research teams become permanent NASS records. This information is used by NHTSA for a variety of purposes, including:

- Assessment of the overall state of traffic safety, and identification of existing and potential traffic safety problems.
- Obtaining detailed data on the crash performance of passenger cars, light trucks, vans, and utility vehicles.
- Evaluation of vehicle safety systems and designs.
- Increasing knowledge about the nature of crash injuries, as well as the relationship between the type and seriousness of a crash and the resultant injuries.
- Assessment of the effectiveness of motor vehicle and traffic safety program standards.
- Evaluation of the effect of societal changes, such as increased traffic flow and increased large truck traffic.

Content copied from NHTSA website,
http://www.nhtsa.gov/Data/National+Automotive+Sampling+System+(NASS)/NASS+Crashworthiness+Data+System on December 10, 2014.