Fall 2011 | **For Members and Friends of the NV ASA**

SILVER STATE-ISTICS

NEVADA Chapter of American Statistical Association

Nevada Chapter News Letter | **Vol. 8 No. 2**

A VERY MERRY CHRISTMAS AND A HAPPY NEW YEAR!

Seasons Greetings

Read on for information on what our Chapter has been up to and what we are planning for the future.
Please keep an eye on our website for more details on future events.

www.amstat.org/chapters/nevada/

Contents:

⌘ ⌘ ⌘

✍ COMING EVENTS ✍

Look for these events to come in the Spring 2012:

- Statistics Career Fair in the North.
- Workshop for High School AP Statistics teachers in the South.
- Dinner meeting featuring Dr. Jay Shen of the UNLV School of Community Health Sciences.
- 2012 K-12 Poster Competition.
- And more!

PRESIDENT'S MESSAGE

Hokwon A. Cho

For the past two years, I have been fortunate to serve as the president of the Nevada Chapter of the American Statistical Association. Though I am a founding member and have also been a chapter representative, it was a great opportunity for me to look into the autonomous organization and the broader national organization. I also welcome the new members who joined recently and the friends who are interested in our organization.

In the next year, the NV-ASA Chapter will celebrate its 10th Anniversary. As an organic body, we will see how much we have been moving forward. Still we have lots of work and vision to make our chapter more viable and strong – carrying out continuously active chapter programs and increasing membership. These fundamental goals and aims cannot be achieved without the devotion of all of our members. Your initiatives, thoughts, time and efforts are the essential fertilizer to sustain and to grow our chapter.

Above all, I thank you all: members of the chapter, participants, speakers, and many volunteers for the memorable events and programs that we have had. Especially, thank you to the Executive Committee members; Ania Panorska (Northern VP), Deb Stiver (Chapter Rep), Alejandra Livingston (former Secretary), Alicia Chancellor Hanson (Webmaster and President elect), Andy Chance (Southern VP), Kaushik Ghosh (Treasurer), Hong Wei (Secretary) , Anton Westveld (former Chapter Rep), Dave Thiel (Education Committee Chair) and Charles Davis (Past President).

Perhaps you are familiar with the popular Christmas lyric, "It's the most wonderful time of the year." This is really meaningful time for all of us for many aspects including the ending-the-year and the beginning-the-new -year. At this time, I want to look back on all the programs that we have accomplished: membership drive, Poster Competitions, Pizza socials, spring/fall statistics symposia, First World Statistics day exhibition, dinner meetings, and Chapter Annual meetings and so on.

With all these memories, I am stepping down from the position but I still remain as a member (and as the past president to ensure a smooth transition of the leadership for the chapter). I am planning to continue working on some of the events that I could not implement previously like the Workshop for K-12 Math Teachers in Nevada. I also challenge you all to continue to support our traditions: what we have done and what we will do under our new leadership.

I truly pray and wish you all, members and friends, the best of luck, wonderful health and more prosperity in the upcoming year!

Elections 2011

Election results were announced during the recent Annual Meeting. We have a new President, Alicia Chancellor Hansen, and a new Chapter Representative, Debra Stiver, both of whom have long-time associations with NV-ASA. Anna Panorska was elected to continue as Northern Vice President, and Hong Wei was elected to continue as Secretary. Our Constitution provides for a staggered election of officers in order to ensure continuity; in 2012 the Southern Vice President and Treasurer offices will be balloted. The Chapter Representative is a three-year position, consistent with national ASA rules; the others are two-year positions.

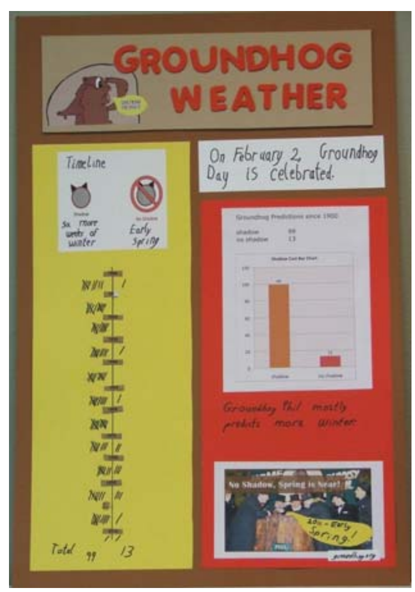
Thanks are due to all of the officers and others who contribute generously of their time and talent!

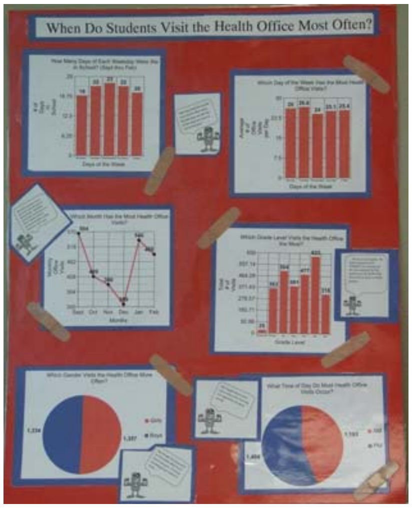
2011 K-12 Poster Competition Update

We are particularly proud of the continued successes of our Nevada K-12 Poster Competition winners in the National Poster Competition. This year Samantha Carpenter and Emily Marcum of Estes McDoniel Elementary School took First Place in the Grades 4-6 category of the national competition with their poster *When Do Students Visit the Health Office Most Often?* Dave Thiel, our Poster Competition leader, is shown here presenting the national award to Emily (left) and Samantha and their advisor, Barbara Chambers.

In addition, Carson McCue of the Alexander Dawson School received an Honorable Mention in the Grades K-3 category for *Groundhog Weather* and Austin Chiu of Hyde Park Middle School received an Honorable Mention in the Grades 7-9 category for *Plants + Sugar = ?*. Their advisors are Betty Aurbach and Denise Romonoski respectively. We observe that our Nevada entrants received three of the nineteen national awards in 2011; no other national winners reside west of Michigan or south of Maryland.

These and the other national winners were featured in the August 2011 issue of *Amstat News*. Here are our winning posters.





All Nevada winning posters may be viewed at www.amstat.org/chapters/nevada/.

NV-ASA Fall Symposium and Annual Meeting 2011

NV-ASA held its Fall Symposium and Annual Meeting Saturday, October 22, 2011 at the University of Nevada, Reno campus. Five speakers presented papers on a variety of topics. Lunch was provided by the UNR Department of Mathematics and Statistics, and a short business meeting completed the day.

Francisco Samaniego of the University of California, Davis, lead the presentations with “A (Very) Short Course on Comparative Statistical Inference”, in which he discussed the Bayesian vs Frequentist dichotomy from the perspective of when or whether the nature of the prior information (or distribution selected) adds to the precision of estimation. This differs from more typical discussions of that dichotomy based on philosophical principle and/or on the nature of the application at hand. When asked for a reference for further discussion, Dr. Samaniego referred to a 1994 JASA paper by Samaniego and Reneau.

The remaining speakers were all from UNR. “Simple Statistics to Explain Genetic Association of Obesity in Nevada” was presented by Karen Schlauch of the Department of Biochemistry and School of Medicine, followed by “Self-similar Trees: An Informal Overview of Theory and Geophysical Applications” given by Ilia Zaliapin of the Department of Mathematics and Statistics. After lunch were heard “Sieve Likelihood Estimation of Partially Linear Poisson Regression with Single-index Model”, presented by Minggen Lu of the School of Community Health Sciences, and “Multivariate Models Connected with Sums and Maxima of Exponential Variables”, presented by Tomasz Kozubowski of the Department of Mathematics and Statistics. Further discussions of the Schlauch, Zaliapin, and Kozubowski presentations are included elsewhere in this Newsletter.

AP Statistics

David Thiel, Coordinator of K–12 Mathematics for the Clark County School District, presented “Statistics Education in the Clark County School District” during the NV-ASA Spring Dinner at Marie Callender’s in Las Vegas March 2, 2011. He provided an essay “Statistics Education in Grades K–12: Then, Now, and in the Future” on the topic, which appeared in our Spring 2011 Newsletter.

More recently, Dave and co-authors have published a paper “AP Statistics: Building Bridges Between High School and College Statistics Education”, which appeared in the August 2011 issue of *The American Statistician*. The article reviews the history of AP (Advanced Placement) Statistics in high schools and the development of the AP Statistics Exam. It compares the content of the AP Statistics curriculum with that of typical entry-level college and university statistics courses. It then addresses current challenges, including the demands placed on high school teacher preparation, adapting those introductory college classes to accommodate students entering college with the background provided by AP Statistics, and identifying and tracking the college careers of incoming students with AP Statistics in their backgrounds.

UNLV Statistics Colloquium/Seminar Series

The UNLV Mathematical Sciences Department maintains an active seminar series, meeting on Fridays at 11:00 a.m.-12:00 p.m. every few weeks. Notices are sent to the NV-ASA email list; for more information go to http://faculty.unlv.edu/cho/statseminar.htm. ❖

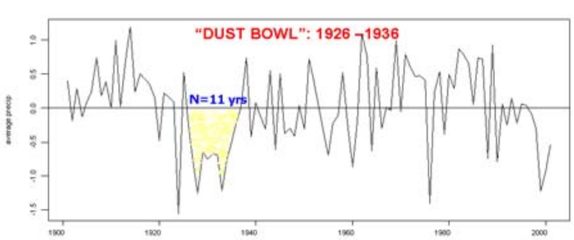
From the Fall Symposium:

Multivariate Models Connected with Sums and Maxima of Exponential Variables

This was presented by Tomasz J Kozubowski of the UNR Department of Mathematics and Statistics; co-authors are Anna K Panorska and Fares Qeadan. The authors present new results about the distributions of sums and maxima of N exponential variables, with N either fixed or having a geometric distribution, as promised in the Symposium announcement.

Their model deals with discrete time series data in which individual observations are classified as positive or negative with respect to a central value; they then deal with "events" which are periods of successive positive (or negative) observations. N is the number of data values in the event, X is the sum of the values, and Y is the maximum. The model is motivated by climate (precipitation, stream flow, etc.) and financial (stock market) studies. Historically only asymptotic results had been available (X_∞ ~ Normal independent of Y_∞ ~ Extreme Value). The new results include explicit joint and marginal pdfs and cdfs along with estimation procedures and more. One of the surprises is that the asymptotic independence is very slow to arrive; even with n = 500 the correlation between X and Y is 0.24. Another somewhat surprising result is that maximum likelihood estimates of parameters of the trivariate distribution of (X, Y, N) do not depend on the value of Y.

Returning to the motivating climatology application, the authors use their model to estimate the probability of a drought longer or larger than the "Dust Bowl" (U.S., 1926-1936) as 0.08, which is far from remote. A copy of the presentation is available on our website.



Self-Similar Trees: An Informal Overview of Theory and Geophysical Applications

This was presented by Ilya Zaliapin of the UNR Department of Mathematics and Statistics; co-authors are Yevgeniy Kovchegov (Oregon State U), Ellen Webb (UNR), and Andrew Hicks (UNR). Dendritic (tree-like) structures occur in many settings; the authors show examples from physiology, botany, hydrology, and even genetics. Applications in genealogy go back at least to Sir Francis Galton's (1873) considerations of the probability of British aristocratic families becoming extinct; hydrogeology applications go back at least to Robert Horton's (1945) investigations of stream and drainage basin formation (see the adjacent image). The presentation reviews basic terminology (such as roots, edges, vertices, and leaves) and concepts (such as Horton-Strahler branching indexing) related to tree structures.

The concept of self-similarity for rooted trees is then introduced. It is related to the idea that branches of the tree have the same structure as the entire tree, in a certain statistical sense. The self-similarity idea is then used in applications to particle dynamics in statistical mechanics (starting with billiard balls) and in studies of earthquake clustering. A copy of the presentation is available on our website.

Simple Statistics to Examine Genetic Associations of Obesity in Nevada

This was presented by Karen Schlauch of the UNR Department of Biochemistry; co-investigators are Doina Kulick and Cynthia Mastick of the UNR School of Medicine. Obesity is an increasing epidemic in Nevada and elsewhere. The authors describe their recent small pilot study aimed at discovering possible associates between obesity and genetics. They present first the study design, involving 20 cases (extremely obese individuals) and 20 controls (lean individuals) with no evidence of typical confounding health factors, carefully matched for gender and age. They describe the genetic assay technique used and the data that result. The statistical tools used are familiar (linear and logistic regression, chi-square tests, and odds ratios). Using FDR-adjusted comparisons (see the accompanying article), they find case-control differences in 58 of some 700,000 Single Nucleotide Polymorphisms (SNPs) examined, controlling the FDR at 5 percent. These include some "surprising" local groupings of SNPs, including one region reported in the literature to be associated with breast, colon, and prostate cancers.

Multiple Comparisons and the False Discovery Rate

By Charles B. Davis

In "Simple Statistics to Examine Genetic Associations of Obesity in Nevada" Karen Schlauch used the acronym "FDR", which was unfamiliar to me. It stands for "False Discovery Rate", a concept introduced by Benjamini and Hochberg (B&H) in 1995 for dealing with the issue of comparing many hypotheses at the same time during a study or experiment. The multiple comparisons (MC) issue arises in many application areas, and has been studied since Dunnett's classic 1955 paper if not earlier. These application areas include medical or drug studies (perhaps a dozen or more main comparisons) and environmental monitoring at regulated facilities (from a few dozen to a thousand or so comparisons), but the magnitude of the MC issue really becomes challenging in genomic studies (often 10K-1M tests, per Karen's presentation).

In the regulatory setting one assumes that all null hypotheses are true (the statistical legal "innocent until proven guilty"). In that setting, if one wants to control the "Family-Wise Error Rate" (FWER) at 10% per year, for example, as suggested in recently released U.S. EPA guidance on groundwater monitoring, AND one would be willing to assume that all of the m_0 hypothesis tests are statistically independent, one could use a per-comparison significance level or error rate (PCER) satisfying $(1 - \text{PCER})$ ^ $m_0 = (1 - 0.1)$. For example, if there are 50 comparisons per year, one would use a per-comparison significance level of .002105; for 1000 comparisons, 0.000105. Alternately, there is the well-known conservative Bonferroni approach of simply using PCER = FWER / m_0; in these two examples the conservative PCERs would become 0.002 and 0.0001 respectively. Note that the Bonferroni approach is very mildly conservative when the tests are actually statistically independent. (The Bonferroni approach would be exact if, somehow, at most one of the null hypotheses could be rejected.) It becomes more conservative if the hypothesis tests are somehow positively correlated. With very low PCERs the statistical power for detecting true alternate hypotheses can be rather low, of course.

But in other studies, such as medical studies, one does not want to assume that all null hypotheses are true, and is very interested in improving the power of tests to detect promising treatments or genetic associations, etc. This is where B&H and their predecessors come in. B&H invented the FDR concept. FDR is the proportion of rejected null hypotheses that are actually true. The idea behind the name is that a rejected null hypothesis is a "discovery". They set up the familiar table. The total number of hypotheses is m, of which m_0 are true null hypotheses. The entries inside the cells are counts, which cannot be observed. What are observed are W and R. FWER = Prob(V > 0). B&H define FDR = E[V / max(R,1)], where E[ ] denotes expected value.

	Accept H_0	Reject H_0	Total
H_0 True	U	V	m_0
H_0 False	T	S	m_1
Total	W	R	m

A main result is that a neat modification of the Bonferroni procedure allows one to control the FDR while at the same time keeping FWER as controlled as in the classical Bonferroni procedure when all null hypotheses are true, but achieving better power when some alternate hypotheses are true. If α is the target FDR and one can assume that the tests are independent (or positively correlated), do the following. Compute p-values for the m tests and sort them $p(1) \le p(2) \le \ldots \le p(m)$; let k be the largest index i for which $p(i) \le (i/m)\ \alpha$; then reject all $H_0(i)$ for $i = 1, \ldots, k$. B&H show that this approach does indeed control FDR at α, and also controls FWER as desired. The gain in power comes from comparing only the lowest p-value with the stringent Bonferroni PCER; if that is rejected, larger p-values are compared with values that are not so low. If one is not willing to assume that the hypothesis tests are independent or positively correlated, the same control can be achieved by taking k to be largest index i for which $p(i) \le (i / (m + 1 - i))\ \alpha$.

References:

Y. Benjamini and Y.Hochberg (1995), "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing", JRSS-B, v. 57, 289-300.

C.W. Dunnett (1955), "A Multiple Comparisons Procedure for Comparing Several Treatments with a Control", JASA, v. 50, 1096-1121.

Book Review: *Neyman From Life* by Constance Reid (1982)

Anna K. Panorska

This enchanting biography of Jerzy Neyman (1894-1981) reads like history of modern statistics. Reid met with Neyman (Professor, Department of Statistics, UC Berkeley) on several afternoons in 1978-79 to talk about his life and work. She also interviewed Neyman's friends and collaborators. The result is an interesting book on the life and times of one of the most prominent statisticians of the 20th century. The book provides a lively account of many major developments in statistics. We learn about the development of the theory of estimation and hypothesis testing, including how the Neyman-Pearson Lemma (with Egon Pearson), a fundamental result in statistical decision theory, came to life, and how maximum likelihood estimation became a standard method in statistics. The book also chronicles the establishment, by Neyman, of the Department of Statistics at UC Berkeley.

Reid draws from memories and reminiscences of prominent statisticians, many of them students of Neyman, such as Peter Bickel, David Blackwell, Joseph Hodges, Lucien LeCam, Erich Lehman, and Elizabeth Scott, among others. We find out about Neyman's theoretical and applied work in many areas of science; in particular, the establishment of the Berkeley Statistical Laboratory. The book is fascinating reading, reflecting the times when some of the main ideas of statistics were being developed. It is accessible to anyone; no knowledge of mathematics or statistics is required. I read it almost without stopping, and recommend it as a great read to anyone interested in the development of modern science.

Baseball Fans – Book/Movie Review Needed

Wanted for the next Newsletter: a review of *Moneyball*, the 2003 book (by Michael Lewis) and/or 2011 movie (starring Brad Pitt), the story of how statistics saved the 2002 Oakland Athletics. The book also provides an introduction to the field of SABeRmetrics (from the Society for American Baseball Research). Will some statistically minded sports fan please "step up to the plate"?

The NV-ASA Mission

From our Constitution:

"The purpose of this Chapter is to be a state and local area leader in promoting all aspects of sound statistical science and practice, including the advancement of statistical education, professional practice, and research."

"The Chapter shall sponsor educational, professional and scientific activities that benefit professional statisticians, students, and users of statistics. The Chapter shall promote effective unified action amongst all groups having an interest in or concern with statistics and probability. The Chapter shall foster communication between the statistical community, other scientific organizations, and the media for the enlightenment of the general public and enhancement of the general welfare. The Chapter shall promote the proper application of statistics in all fields, shall work for the improvement of statistical education at all levels, and shall seek opportunities to advance the statistics profession and professional work opportunities." ♣

Joining NV-ASA!

Only a minority of the people who receive this newsletter are members of the Nevada Chapter of the American Statistical Association (MV-ASA). Dues are nominal: at most $20 per year, $10 a year for members of the national ASA, and $2 for student members of the national ASA. There are two ways to join. You can join NV-ASA when you renew your national ASA membership (or join for the first time); this can be done on-line at www.amstat.org. Otherwise, whether a national ASA member or not, you can join through PayPal on our website www.amstat.org/chapters/nevada/ or by contacting our Treasurer Kaushik Ghosh. Any way you do it, please obtain an information form from our website, complete it, and send it to Kaushik at the address listed on the form.

Why should you join? NV-ASA events provide opportunities for networking and contact with other statisticians working in a wide variety of areas in Nevada. But in addition to that, a major reason is that your dues support the outreach activities of the NV-ASA including the K-12 Poster Competition and Career Days. Our financial needs are not great, so long as we all pitch in our modest amounts. ⌘

San Diego JSM 2012

July 28 - August 2, 2012
San Diego, California,
San Diego Convention Center

NV-Chapter Current Officers

President: Hokwon A. Cho (through 2011)
Alicia Chancellor Hansen (2012-2013)
Past President: Charles B. Davis (through 2011)
Hokwon A. Cho (2012-2013)
Southern Vice President: Andy Chance (through 2012)
Northern Vice President: Anna Panorska (through 2013)
Secretary: Hong Wei (through 2013)
Treasurer: Kaushik Ghosh (through 2012)
Chapter Representative: Debra Stiver (through 2013)
Education Committee Co-Chair: David W. Thiel
Webmaster: Alicia Chancellor Hansen

Editor:
Charles B. Davis

Typesetting:
Hokwon A. Cho

Silver State-istics welcomes news items and letters from members and friends of the NV-ASA on matters of interest to the Chapter and the profession. Manuscript or items can be sent as a Microsoft Word document, PDF, or within an e-mail.
Silver State-istics is published by the Nevada Chapter of the American Statistical Association.

For contact information, please see our website

www.amstat.org/chapters/nevada/