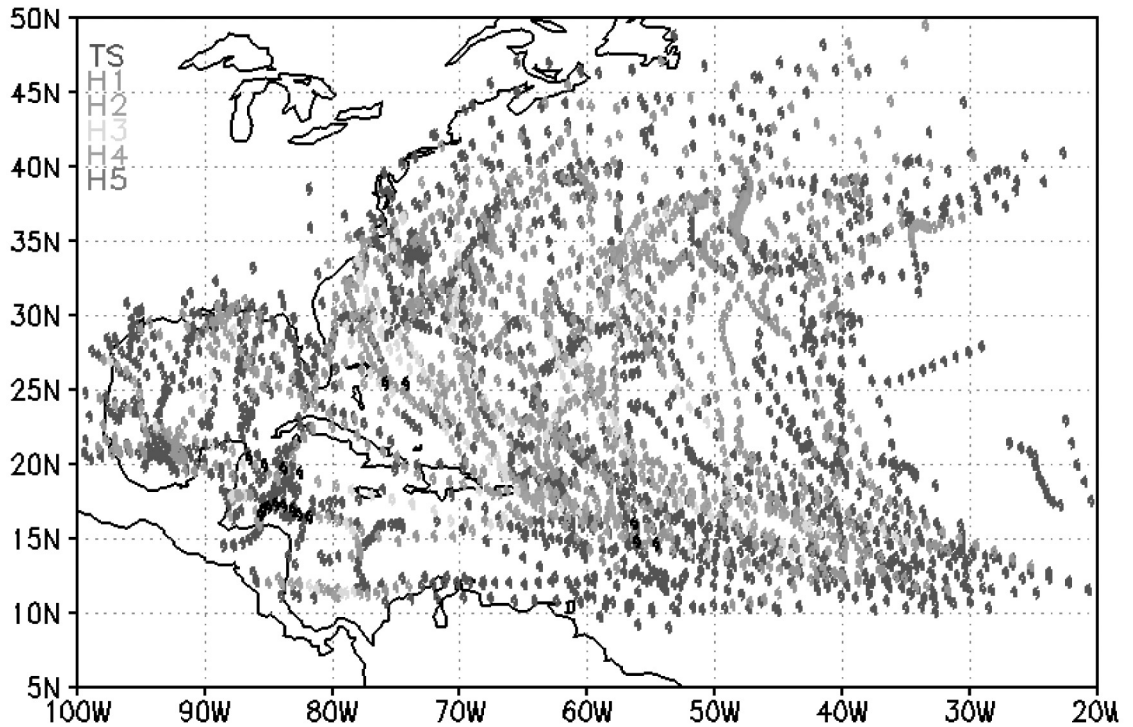


The Statistics of

# HURRICANES



**Interview with Alan Agresti**  
**When Direction Vanishes**  
**Statistics Projects with a Future**  
**Can Baseball Teams Buy Success?**

Non Profit Org-  
U.S. Postage  
PAID  
Permit No. 361  
Alexandria, VA

# Access Professional Resources with the American Statistical Association

## MEMBERS RECEIVE:

A SUBSCRIPTION to *Amstat News*, ASA's monthly membership magazine, full of upcoming events and job opportunities.

NETWORKING through ASA's regional Chapters and special interest Sections.

MEMBERS ONLY FEATURES of the ASA JobWeb at <http://jobs.amstat.org/> including "Post Your Resume" and "Notify Me!"

OPPORTUNITIES to expand your career horizons with ASA's career Placement Service at the Joint Statistical Meetings.

DISCOUNTED registration fees for the annual Joint Statistical Meetings and Continuing Education Courses.

## Become a MEMBER of ASA join at [www.amstat.org/join](http://www.amstat.org/join)

All membership categories include a subscription to *Amstat News*. Student members also receive *STATS: The Magazine for Students of Statistics*. All members receive discounts on ASA publications, meetings, and educational programs. Information regarding your benefits, including how to subscribe to additional publications or register for ASA meetings will be included in your new member package. Allow 2-4 weeks for processing of your membership. If you have any questions, please contact customer service at **1-888-231-3473** or visit our website [www.amstat.org/join](http://www.amstat.org/join).

**YES**, I want to join now (select one):

Regular Membership.....\$85     Senior Membership.....\$25     Student Membership.....first year now ONLY \$10

Maybe, please send me an **ASA Membership Guide** to review.     No thank you, but continue to send information about ASA.

**Member Information** (circle one)    DR.    MR.    MS.    MISS    MRS.

**Name** First: \_\_\_\_\_ Middle \_\_\_\_\_ Last \_\_\_\_\_

Informal Name: \_\_\_\_\_ Organization \_\_\_\_\_

Mailing Address \_\_\_\_\_

City \_\_\_\_\_ State/Province \_\_\_\_\_ Zip/Postal Code \_\_\_\_\_ Country \_\_\_\_\_

Email Address \_\_\_\_\_ Office phone \_\_\_\_\_ Home Phone \_\_\_\_\_

**Method of Payment**     Payment for \$ \_\_\_\_\_ is enclosed (make payable to ASA in U.S. funds drawn on a U.S. bank)

Charge my credit card for \$ \_\_\_\_\_     American Express     Mastercard     Visa

card # \_\_\_\_\_ expiration \_\_\_\_/\_\_\_\_

signature \_\_\_\_\_     bill me

visit us at  
[www.amstat.org](http://www.amstat.org)  
or call customer service at  
1-888-231-3473



# The Magazine for Students of Statistics

## Winter 2004 • Number 39

### Editors

**Beth L. Chance**

email:  
bchance@calpoly.edu

Department of Statistics  
California Polytechnic State University  
San Luis Obispo, CA 93407

**Allan J. Rossman**

email:  
arossman@calpoly.edu

Department of Statistics  
California Polytechnic State University  
San Luis Obispo, CA 93407

### Editorial Board

**Patti B. Collings**

email:  
collingp@byu.edu

Department of Statistics  
Brigham Young University  
Provo, UT 84602

**E. Jacquelin Dietz**

email:  
dietz@stat.ncsu.edu

Department of Statistics  
North Carolina State University  
Raleigh, NC 27695-8203

**David Fluharty**

email:  
fluharty\_david@hotmail.com

Continental Teves  
One Continental Drive  
Auburn Hills, MI 48326

**Robin Lock**

email:  
rlock@stlawu.edu

Department of Math, CS, and Stat  
Saint Lawrence University  
Canton, NY 13617

**Chris Olsen**

email:  
colsen@cr.k12.ia.us

Department of Mathematics  
George Washington High School  
Cedar Rapids, IA 53403

**Josh Tabor**

email:  
jtabor@hlpusd.K12.ca.us

Glen A. Wilson High School  
16455 Wedgeworth Drive  
Hacienda Heights, CA 91745-3045

### Production

**Jennifer Campanile**

email:  
jennifer@amstat.org

American Statistical Association  
1429 Duke Street  
Alexandria, VA 22314-3415

**Michael Campanile**

email:  
michaelc@amstat.org

American Statistical Association  
1429 Duke Street  
Alexandria, VA 22314-3415

**Megan Murphy**

email:  
megan@amstat.org

American Statistical Association  
1429 Duke Street  
Alexandria, VA 22314-3415

*STATS: The Magazine for Students of Statistics* (ISSN 1053-8607) is published three times a year, in the winter, spring, and fall, by the American Statistical Association, 1429 Duke St., Alexandria, Va 22314-3415 USA; (703) 684-1221; fax: (703) 684-2036; Web site: [www.amstat.org](http://www.amstat.org).

*STATS* is published for beginning statisticians, including high school, undergraduate, and graduate students who have a special interest in statistics, and is distributed to student members of the ASA as part of the annual dues. Subscription rates for others: \$13.00 a year to members; \$20.00 a year to nonmembers.

Ideas for feature articles and material for departments should be sent to the Editors; addresses of the Editors and Editorial Board are listed above.

Requests for membership information, advertising rates and deadlines, subscriptions, and general correspondence should be addressed to *STATS* at the ASA office.

Copyright © 2004 American Statistical Association.

### Features

- 3 The Statistics of Hurricanes  
*Madhuri S. Mulekar and Sytske Kimball*
- 10 Interview with Alan Agresti  
*Jackie Dietz*
- 15 Statistics Projects with a Future  
*George Kesling and Wayne Fairburn*
- 17 Play ASA Stat Bowl in Toronto!  
*Mark Payton*
- 29 Advice from the 2003 Stat Bowl Champion  
*Wesley Eddings*

### Departments

- 2 **Editors' Column**
- 18 **Student Project**  
When Direction Vanishes: Walking Straight or in Circles  
*Andrea L. Axtell*
- 22 **The Statistical Sports Fan**  
Can Major League Baseball Teams Buy Success?  
*Robin Lock*
- 26 **AP Statistics**  
A Simulation Alternative to the Two Sample *t*-test  
*Josh Tabor*
- 28  **$\mu$ -sings: Statistics in the Media**  
Review of *Life of Sir Francis Galton: From African Exploration to the Birth of Eugenics*, by Nicholas Wright Gillham  
*Chris Olsen*

Hurricanes are very destructive natural forces, but they also display a power and beauty that fascinate meteorologists and other scientists. In this issue's lead article, Madhuri Mulekar and Sytske Kimball reveal that hurricanes also provide plenty of what statisticians find irresistible: data. Professors Mulekar and Kimball analyze a dozen variables on thousands of hurricanes in an effort to better understand the structure and behavior of these storms. They point out that statisticians have an important role to play in this research, which they hope will lead to better predictions about hurricane activity and, ultimately, to saved lives.

Jackie Dietz turns her attention to Alan Agresti as an interview subject in this issue. Professor Agresti is probably best known for his textbooks on categorical data analysis, particularly with applications to the social sciences. He describes how he came to study statistics during the turbulent Vietnam War era, how he changed his research interests dramatically after finishing his dissertation, and how he believes that statisticians entering academia today face higher expectations than his generation did. Professor Agresti also offers advice for students considering a career in statistics.

George Kesling and Wayne Fairburn offer an article in which they suggest that statistics students consider transforming their class research projects into presentations at professional conferences. They provide practical advice and point out the benefits of such professional activity. We are very sad to note that Professor Kesling died shortly after preparing this article for *STATS*; we hope that its publication will honor his memory and his dedication to his students.

Mark Payton recaps the Stat Bowl competition held at the Joint Statistical Meetings in San Francisco last August, and he invites students to participate in the upcoming summer's contest in Toronto. Last year's champion Wesley Eddings contributes some encouragement and advice for this year's participants. Stat Bowl provides a way to challenge yourself, both intellectually and trivially at the same time! ASA offers travel funding for student participants in the competition, which provides an opportunity to meet



**Beth Chance and Allan Rossman**

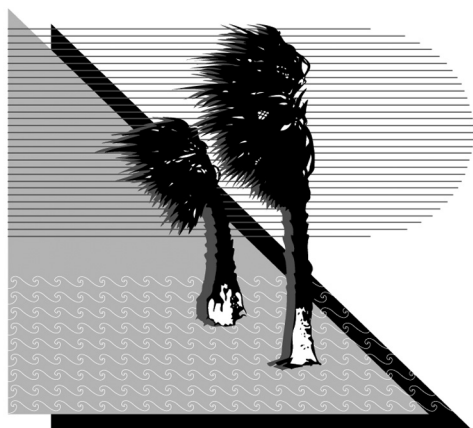
other students at the conference. Please read about the Stat Bowl, try your luck with the sample questions, and consider applying to participate.

In last spring's issue we featured an article about a student project conducted by the 2002 winner of the ASA Award at the International Science Fair. In this issue we are very pleased to present an article by the 2003 winner, Andrea Axtell. Andrea's project related to the question of whether people who are lost in the woods (and, by extension, characters like Pooh and Piglet as well) tend to walk in circles. Andrea explored this by blindfolding subjects and asking them to walk on a football field from one goal line to the other. She also analyzed the relationships between subjects' height, gender, and handedness and their ability to walk straight.

Robin Lock is not only the statistical sports fan for *STATS*—he is also a diehard Yankees fan bitterly disappointed by the Florida Marlins' World Series victory in 2003. In this issue's column he considers the question of whether a baseball team's payroll is associated with its success on the field. Robin demonstrates how several different statistical techniques can be applied to this issue, depending on how one defines the variables of interest.

In this issue's AP Statistics column, Josh Tabor writes about a fundamental statistical problem: comparing measurements between two groups. This sounds like a situation that calls for a  $t$ -test, but Josh tackles the problem through simulation. He also makes the case that students can develop a firm intuitive understanding of statistical significance and  $p$ -values through this type of analysis.

Chris Olsen considers a genre that he usually avoids in this issue's  $\mu$ -sing. He recommends a biography of Francis Galton that describes Galton's early career as an African explorer and geographer before he invented the techniques of correlation and regression that so many scientists and statisticians employ to this day. Chris also comments on Galton's role as pioneer of the controversial subject of eugenics and reveals a Fisher/Pearson-like feud that Galton found himself entangled in. ■



# The Statistics of Hurricanes

Hurricanes are beautiful natural phenomena, but highly destructive and dangerous to life and property. In order to provide better forecasts of their tracks and damage potential, a large amount of meteorological research into hurricane behavior takes place across the globe. This research consists of a combination of statistical analysis of real data, theoretical derivations using mathematical equations and physical laws, and computer simulations. Via remote sensing and flights into hurricanes, a large amount of hurricane data is collected every season. Some of these data span many decades and include a wide range of storms of different intensities and sizes. Such large datasets are usually analyzed using statistical methods. By looking at large datasets that span years and decades, trends in hurricane formation, motion, structure, and intensity can be isolated. Working in collaboration with meteorologists and other scientists, statisticians can play an important role in helping to understand hurricane behavior and in predicting storm intensities and tracks that can ultimately lead to preventing the loss of life and property.

## Background

What exactly is a hurricane? Hurricanes belong to a group of weather phenomena known as tropical cyclones. A “cyclone” is a low pressure disturbance with winds blowing in an anti-clockwise direction around the center in the northern hemisphere. A tropical cyclone (TC) has several additional features: it is warm-cored, it forms over warm, tropical or sub-tropical waters, its maximum winds occur near the surface, and the main energy source comes from latent heat release in strong

*Madhuri S. Mulekar (mmulekar@jaguar1.usouthal.edu) is a professor of statistics in the Department of Mathematics and Statistics at the University of South Alabama. Her major research areas are selection and ranking procedures and sequential testing procedures. She also works with meteorologists on characterization of hurricanes and with physicians on designing experiments and analyzing clinical data.*

*Sytske Kimball (skimball@usouthal.edu) is an assistant professor of meteorology in the Department of Earth Sciences at the University of South Alabama. Her research focuses on hurricane structure and intensification. She has been working on hurricanes for 10 years, both in Australia and the United States.*



**Madhuri S. Mulekar and Sytske Kimball**

thunderstorm clouds. The term “tropical cyclone” is a collective name for storms with such features and includes storms of all intensities: tropical depressions (TD), tropical storms (TS), and hurricanes (also known as “typhoons” in the North Pacific Ocean).

Hurricanes are classified into different categories based on their intensity (maximum low-level windspeed or minimum surface pressure). Meteorologists use the Saffir-Simpson (S-S) scale (Simpson, 1974) to classify hurricanes (see Table 1), and emergency managers use this classification to communicate the potential threat of an approaching storm to the general public. For example, Hurricane Isabel (2003) was a category 5 hurricane at the time of the image in Figure 1, the most damaging category. Based on a post-analysis of wind damage data and recently improved observations of low-level wind profiles in hurricanes, Hurricane Andrew (1992) was upgraded from a category 4 to a category 5 hurricane. This severe hurricane caused \$26.5 billion worth of damage and 65 deaths in the Miami area and in southeast Louisiana.

Category	Windspeed range (ms <sup>-1</sup> )	PSMIN range (hPa)	Storm Surge (m)
Tropical storm	17-33	—	—
1	34-42	>980	1.0-1.7
2	43-49	979-965	1.8-2.6
3	50-58	964-945	2.7-3.8
4	59-69	944-920	3.9-5.6
5	>70	<920	>5.7

In this article, 15 years of data are used to study several hurricane characteristics. The dataset is known as the Extended Best Track (EBT) dataset (Pennington, DeMaria, and Williams, 2000) and contains North Atlantic tropical cyclone data from 1988 to 2002. Measurements were taken for all storms at six-hour

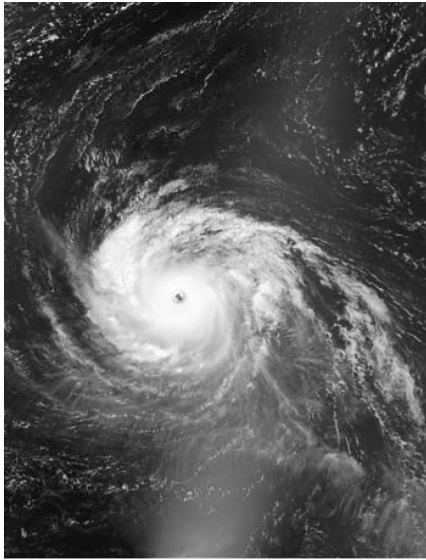


Figure 1: Visible satellite image of Hurricane Isabel, 14:15 GMT, September 11, 2003

intervals from the tropical depression to the dissipation stage. The data file originally contained 5,638 records of six-hourly observations; data on tropical depressions and extra-tropical storms were removed. After performing consistency checks, 3,493 records for 172 storms remained. Along with the storm-name, -year, -month, -day, and -hour, the following variables were recorded at six-hourly intervals, for each storm:

- PSMIN: Minimum sea-level pressure, measured in hecto-Pascals (hPa).
- VMAX: Maximum low-level windspeed, measured in meters per second ( $ms^{-1}$ ).
- LAT: North-south location of storm measured in degrees latitude ( $^{\circ}N$ ).
- LON: East-west location of storm measured in degrees longitude ( $^{\circ}W$ ).
- REYE: Radius of the eye of the storm, measured in kilometers (km).
- RMW: Radius of VMAX, i.e., the distance between the center of the storm and the location of the maximum winds, measured in kilometers (km).
- R33: Radius of  $32.9 ms^{-1}$  winds, i.e., the distance between the storm center and the location of the  $32.9 ms^{-1}$  winds, measured in kilometers (km).
- R26: Radius of  $25.7 ms^{-1}$  winds, i.e., the distance between the storm center and the location of the  $25.7 ms^{-1}$  winds, measured in kilometers (km).
- R17: Radius of  $17.5 ms^{-1}$  winds, i.e., the distance between the storm center and the location of the  $17.5 ms^{-1}$  winds, measured in kilometers (km).
- ROCI: Radius of outermost closed isobar (an isobar is a line of constant pressure), measured in kilometers (km).

- POI: Pressure of the outermost closed isobar, measured in hecto-Pascals (hPa).

## Relationship between minimum sea-level pressure and maximum windspeed

Hurricanes have several distinct features that are easily recognizable. Figure 1 shows a satellite image of Hurricane Isabel taken at 14:15 Greenwich Mean Time (GMT) on September 11, 2003. Very noticeable in this hurricane is the relatively cloud-free eye in the center of the storm. Surrounding the eye is a dense and tall wall of clouds that extends from close to the surface to the top of the troposphere (about 15 km into the atmosphere). Outward from this “eyewall” are so-called “spiral rainbands.” Both the eyewall and the rainbands consist of vigorous thunderstorms. In the eye and eyewall (or the core of the storm) at higher levels, the temperature in a TC can be around  $10^{\circ}C$  higher than the surroundings. This warm core causes the surface pressure in the eye to be lower than anywhere else. This low pressure causes a radial pressure gradient from the center outward. Air flows from high to low pressure and, hence, air from outside the storm rushes in toward the eye at low levels, causing the characteristic damaging winds that accompany hurricanes. The rotation of the winds around the storm center is caused by the spinning of the Earth around its axis. In other words, the minimum sea-level pressure (PSMIN) and the maximum low-level wind speed (VMAX) in a hurricane are closely related.

Figure 2 shows a very strong negative relation ( $r = -0.9488$ ,  $p < 0.0001$ ) between VMAX and PSMIN. For storms in the EBT dataset, the relation between VMAX and PSMIN was estimated to be  $VMAX = 642.24 - 0.62 PSMIN$ . This tells us that, for every 1 hPa decrease in minimum surface pressure, the maximum low-level windspeed increases by an average of  $0.62 ms^{-1}$ . Due to this very high correlation between VMAX and PSMIN, the intensity of a tropical cyclone can be measured in terms of either the maximum low-level windspeed or the minimum sea-level pressure.

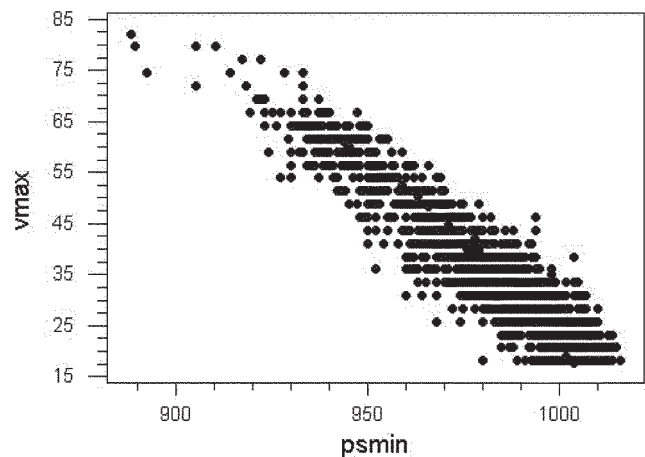


Figure 2: Relation between maximum low level winds and minimum surface pressure

## Size of Hurricanes

As seen from Figure 2, VMAX and PSMIN are strongly correlated, but this relation does not specify how far away from the center the maximum wind occurs. Figure 3 shows a radial cross-section of the low-level windspeed in a hypothetical hurricane. In the center of the storm, conditions are calm. Moving outward from the storm center, the windspeed increases rapidly to its maximum (VMAX). The distance between the location of the maximum winds and the storm center is known as the radius of maximum winds (RMW). Outward from the RMW the windspeed drops off gradually and crosses three key values: (a)  $32.9 \text{ ms}^{-1}$ , the cut-off between hurricane- and tropical storm-force winds; (b)  $25.7 \text{ ms}^{-1}$ , not a cut-off value, but measured traditionally; (c)  $17.5 \text{ ms}^{-1}$ , the cut-off between tropical storm- and tropical depression-force winds. These radial distances are known as the radii of  $32.9$ ,  $25.7$ , and  $17.5 \text{ ms}^{-1}$  winds respectively, or R33, R26, and R17 for short. The maximum winds are located somewhere within the eyewall. The radius of the eye (REYE) indicates where the eyewall begins (Figure 3). The radial distances shown in Figure 3 are known as size parameters and give a sense of the horizontal spread of the hurricane.

The size parameters of all storms in the EBT dataset show a large range in size. Figure 4 shows a distribution of REYE. One observation is identified as an outlier. This is an eye radius of  $50.9 \text{ km}$  for Hurricane Georges (1998), the largest on record in 15 years. The right-skewed nature of the distribution of REYE indicates that most storms have comparatively small eyes. The boxplot indicates that the sizes of these eyes are spread over a relatively short range. Only a few large eyes are observed, but their sizes show considerable spread.

As seen from Figures 5 and 6, the size parameters vary with storm intensity (measured in terms of S-S category). The abbreviations TS and H1 through H5 are used for tropical storms and hurricanes of each of the five S-S categories, respectively. The median REYE (Figure 5) is seen to increase with storm intensity for tropical storms and weak hurricanes (H1 and H2). The most intense hurricanes (H5) tend to have smaller eyes and a shorter spread of possible eye sizes. For most S-S categories, except H2 and H4 storms, the REYE distributions are right skewed. In other words, most of these storms tend to have smaller eyes; only a few have large eyes. Mood's median test (Conover, 1999) was applied to compare the median eye radii of different categories. The result is significant at the 5% level, indicating that the median eye radii are different for at least two categories ( $p < 0.001$ ). Interestingly, both the weakest (TS) and the most intense (H5) tropical cyclones never obtain eyes larger than  $30 \text{ km}$ , whereas the other categories do.

Not all storms have eyes, and higher percentages of eye observations occur in more intense S-S categories (Table 2). In other words, the likelihood of a storm

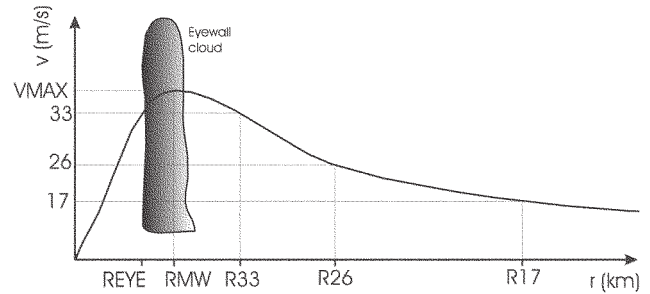


Figure 3: Tropical cyclone wind profile with size parameters

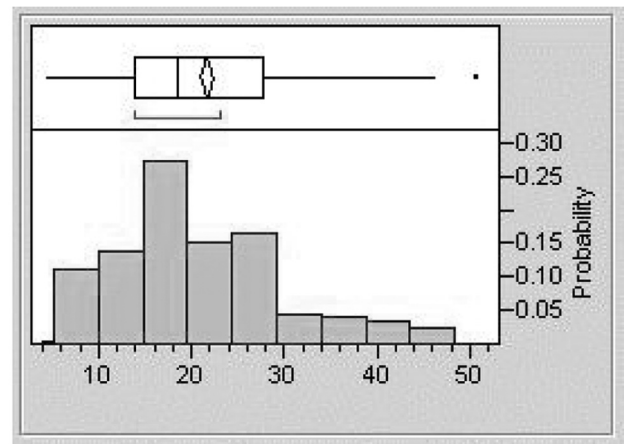


Figure 4: Distribution of eye radius

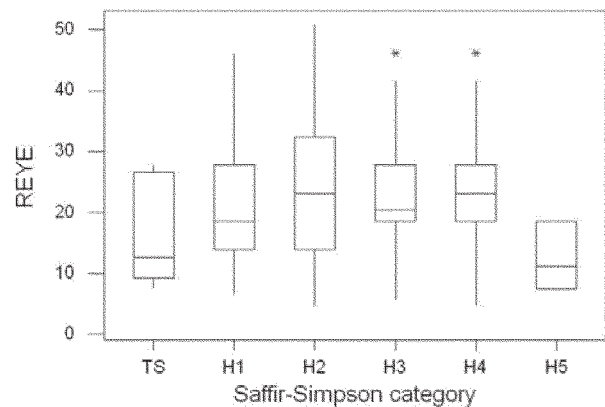


Figure 5: Distribution of eye radius by storm category

Table 2: Number and duration of Atlantic tropical cyclones and percentage of tropical cyclones with eyes from 1988–2002							
	TS	H1	H2	H3	H4	H5	ALL
Number of storms	172.0	97.0	59.0	38.0	24.0	7.0	172.0
% of storms with an eye	5.2	51.5	72.9	81.6	91.7	100.0	38.4
Number of six-hourly obs.	1,912.0	913.0	315.0	177.0	156.0	20.0	3,493.0
Average duration (h)	66.7	56.5	32.0	27.9	39.0	17.0	121.8

having a well-defined eye increases with the storm intensity. The formation of an eye indicates that a TC has reached a high level of organization, and is therefore usually a characteristic of more intense systems (Weatherford and Gray, 1988).

The median RMW is observed to decrease with increasing storm intensity (Figure 6a) from 55.5 km at the TS stage to 27.8 km at the H5 stage, a reduction of 50%. A decreasing spread of RMWs with increasing S-S category can also be seen. Mood's median test confirms that the differences in median RMWs are significant for at least two storm categories ( $p < 0.001$ ). This RMW contraction is consistent with Willoughby's (1998) theory of eyewall contraction of intensifying storms.

Figure 6b shows the variation of R33 with storm category [by definition, tropical storms do not have a radius of  $32.9 \text{ ms}^{-1}$  winds (see Table 1)]. It shows that weak hurricanes have a comparatively smaller median R33, and intense hurricanes tend to have a comparatively larger median R33. This means that in the latter case, winds in excess of  $32.9 \text{ ms}^{-1}$  extend over a larger area. Figure 6b also shows little variation in the medians of R33 for intense hurricanes (H3, H4, and H5), i.e., whereas the maximum windspeed increases substantially with increasing S-S category (Table 1), the aerial coverage of winds greater than  $32.9 \text{ ms}^{-1}$  changes very little.

Apart from H4s, a steady increase in R26 measurements with hurricane category can be seen in Figure 6c. In other words, outside the core (see Figure 3), the aerial coverage of destructive winds increases with increasing intensity. R17 is often used as a measure of overall hurricane size and Figure 6d shows that hurricanes tend to increase in size with increasing category from TS to H3. Very intense hurricanes (H4 and H5) tend to be somewhat smaller than H3s but larger than weak hurricanes and TSs. These differences in size parameters by S-S category are summarized in Figure 7, where the median size parameters from the boxplots and the corresponding windspeeds are plotted for TS, H1, H3, and H4s. By connecting the individual points, a wind profile is drawn. Tropical storms have a relatively flat wind profile as shown in Figure 7a. As tropical storms intensify to H1s, the inner core (0–about 111 km radius) winds increase in magnitude, while the

horizontal extent of destructive (in excess of  $25.7 \text{ ms}^{-1}$ ) winds expands and the size (R17) increases. As further intensification takes place (Figure 7b) the RMW shrinks while the area of destructive winds increases further. The difference between H3 and H4 profiles is interesting: while the maximum winds increase, the aerial extent of damaging winds decreases. More variability is observed in the size of the outer wind parameters (R17 and R26) than the inner core parameters (RMW and R33).

## Location of Hurricanes

Along with the study of the wind structure and the size characteristics of storms, meteorologists are interested in the spatial variability of hurricanes throughout the Atlantic basin and their temporal variability during the hurricane season. Figure 8 shows the distribution of the origin of EBT storms as a function of latitude and longitude. Hurricanes form mostly between  $10^\circ$  and  $20^\circ\text{N}$  (see Figure 8a and map on cover) but the longitude of formation is highly variable (Figure 8b). The longitude of the location where a storm forms varies largely depending on the month of the hurricane season. Table 3 shows the median latitude and longitude of formation for each month. Early- and late-season storms form further west, whereas midseason storms form further east. Once formed, tropical cyclones tend to travel north- and west-ward. Figure 9 (see pg. 9) shows the distribution of latitude for intensifying and weakening hurricanes. Intensifying storms show a slight preference to occur south of  $20^\circ\text{N}$ , while dissipating storms occur further north where conditions become more unfavorable. Unfavorable conditions include

Table 3: Median latitude and median longitude of tropical cyclone formation stratified by month						
	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.
Latitude ( $^\circ\text{N}$ )	26.9	22.4	14.3	14.9	18.9	21.7
Longitude ( $^\circ\text{W}$ )	84.0	71.0	45.9	55.2	65.3	71.4



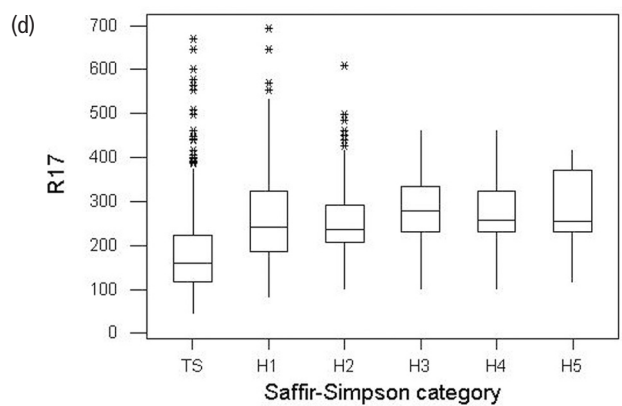
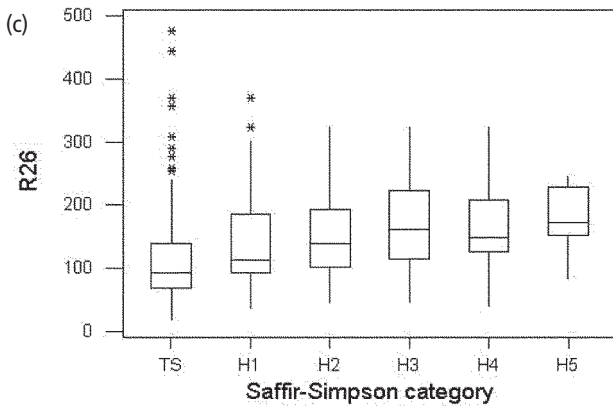
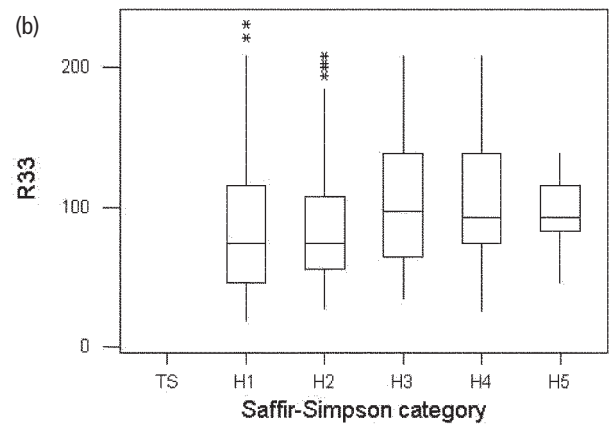
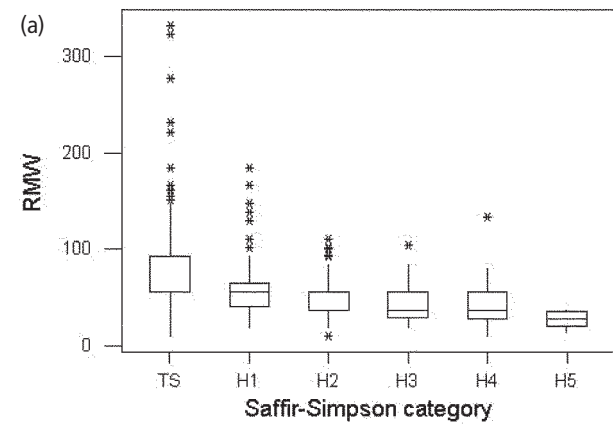


Figure 6: Distribution of size parameters by storm category (a) RMW, (b) R33, (c) R26, (d) R17.

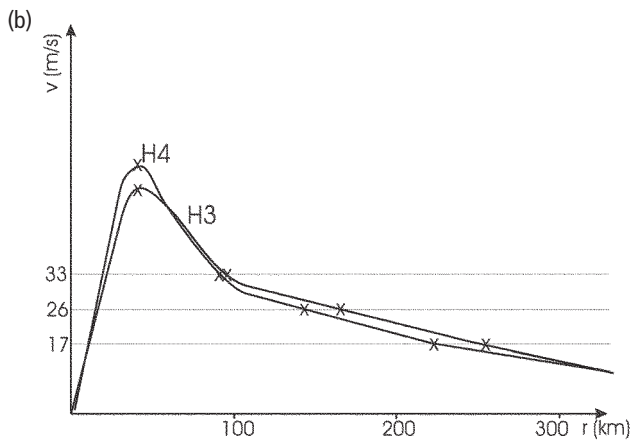
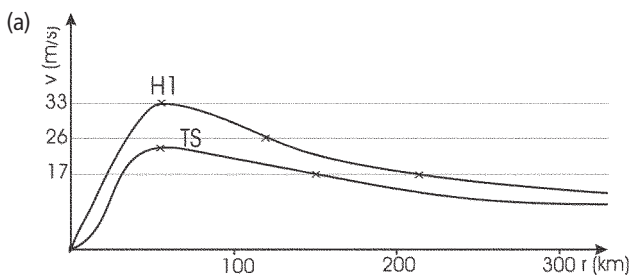


Figure 7: Wind profiles of four tropical cyclone S-S categories; a) TS and H1, b) H3 and H4.

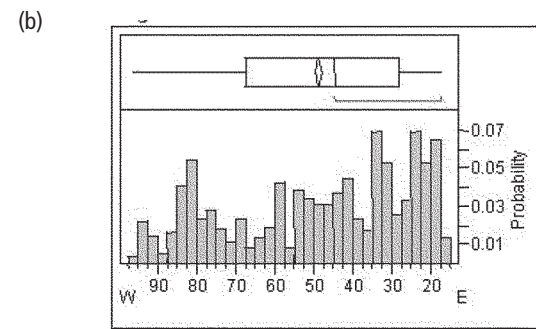
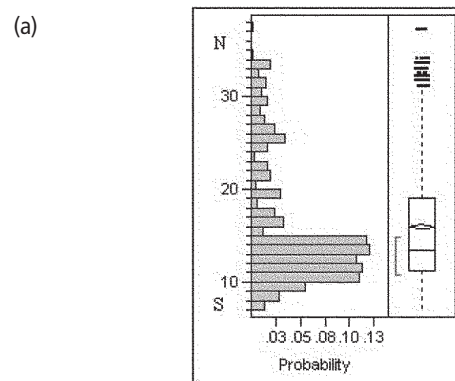


Figure 8: Distribution of the location of origin for storms during 1988-2002; (a) Latitude of origin, (b) longitude of origin.

cooler sea surface temperatures and increased vertical wind shear (windspeed and wind direction changing with height).

## The Life Cycle of Hurricanes

Most hurricanes start out as tropical depressions, and then intensify to TSs and to the various S-S hurricane categories. After obtaining their maximum intensity, they weaken and descend through the various S-S categories. Intensity fluctuations may occur several times during a storm's lifetime. For each storm, data were recorded every six hours during its lifetime. Multiplying the total number of six-hourly observations over 15 years for each Julian day by six converts the total number of six-hourly observations to the total number of "TC hours." TC hours measure the amount of time a storm spends in each TC category. Since the total number of TC hours per Julian day is highly variable, a time series plot of TC hours versus month of the year would be very noisy. A 13-day moving average was used to reduce these jitters (high-frequency fluctuations) in the data and smooth the plot. A period of 13 days was chosen based on previous work by Landsea (1993). Figure 10 shows the 13-day moving average plot of the total number of TC hours per Julian day, stratified by storm category, for Atlantic TCs over 15 years. It shows that the amount of TC hours decreases with increasing S-S category. More tropical storm hours are observed per season than any other category. This could indicate that (a) many TCs never intensify beyond the TS stage, and/or (b) TCs that do intensify to hurricanes spend a long time in the TS category.

The number of storms in each S-S category is listed in Table 2 and concurs with explanation (a). The number of six-hourly observations in each storm category are also listed Table 2. Using this information, the average duration of storms in each category was calculated and plotted in Figure 11. This shows a downward trend of average duration with the increasing storm category, which supports explanation (b). So the EBT data (collected for hurricanes over 15 years) supports both explanations for the high occurrence of TS hours in Figure 10. In other words, very intense hurricanes do not occur often. This indicates that atmospheric conditions are not usually favorable for a TS to develop into an intense hurricane. Furthermore, if such a storm does form, it usually does not last very long, suggesting that the highly organized state of a category 5 hurricane cannot be physically maintained for a very long time.

Consistent with the findings of Landsea (1993), who studied all TCs from 1886 to 1991, very few TCs occur before June 1 and after November 30 (see Figure 10), the dates marking the start and end of the official Atlantic basin hurricane season. Landsea found a distinct peak in Atlantic TC activity in mid-September, but the 15-year EBT dataset shows that weaker systems (TS and H1) peak in late August, earlier than in

Landsea's case. TS and H1s show a secondary peak in late September. The reverse is seen for systems of category H2, a peak in late September with a secondary peak in late August. Intense hurricanes of categories H3, H4 and H5 peak in mid-September, which is consistent with Landsea's observations. Landsea used a 106-year dataset (from 1886 to 1991), compared to only 15 years of data used here. Hence, the dual peaks may be a reflection of a decadal climate variability that did not show up in Landsea's larger multi-decadal dataset.

## Summary

While these beautiful natural phenomena have always fascinated scientists and weather enthusiasts alike, hurricanes are highly destructive and form a serious threat to coastal communities, off-shore interests, and shipping. With the aid of advanced technologies in data recording and data analysis, researchers are learning more about their structure and behavior. Hurricane watchers and researchers alike depend on statisticians to help with collecting, analyzing, and interpreting these data. Ultimately, accurate forecasts of hurricane intensity, size, and track will allow emergency managers to issue precise preparation and evacuation warnings. This will reduce over-warning and hence will save money and lend credibility to the forecasting agencies. Furthermore, this will allow the population adequate time to make preparations to protect their lives and property. However, much remains to be done before this goal can be achieved. With a continuously rising coastal population, the need for top notch forecasting tools, and hence additional hurricane research to develop them, remains critical.

## References

- Conover, W.J. (1999), *Practical Nonparametric Statistics (3rd ed.)*, Wiley & Sons, 218.
- Landsea, C.W. (1993), A Climatology of Intense (or Major) Atlantic Hurricanes. *Monthly Weather Review*, 121, 1703-1713.
- Pennington, J., DeMaria, M., and Williams, K. (2000), Development of a 10-year Atlantic Basin Tropical Cyclone Wind Structure Climatology at [www.bbsr.edu/rpi/research/demaria/readme13.html](http://www.bbsr.edu/rpi/research/demaria/readme13.html).
- Simpson, R.H. (1974), The Hurricane Disaster Potential Scale. *Weatherwise*, 27, 169 and 186.
- Weatherford, C.L., and Gray, W.M. (1988), Typhoon Structure as Revealed by Aircraft Reconnaissance Part II: Structural Variability. *Monthly Weather Review*, 116, 1044-1056.
- Willoughby, H.E. (1998), Tropical Cyclone Eye Thermodynamics. *Monthly Weather Review*, 126, 3053-3067. ■

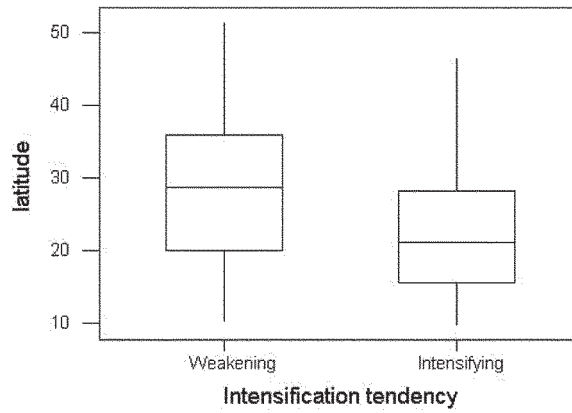


Figure 9: Distribution of latitude by intensification tendency.

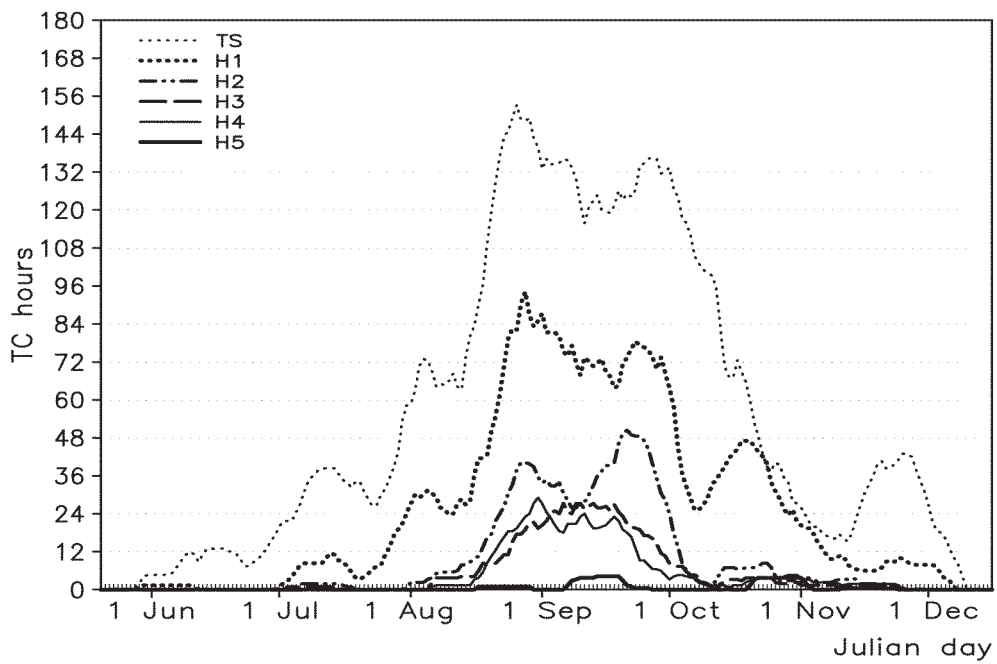


Figure 10: Number of TC hours per Julian day, by Saffir-Simpson category.

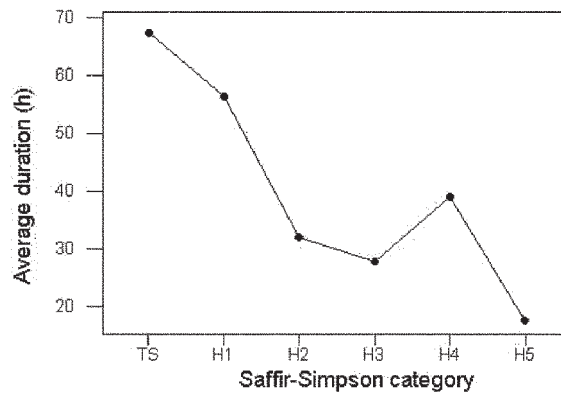


Figure 11: Average duration for storms in each S-S category.

# Interview with Alan Agresti

**A**lan Agresti is Distinguished Professor of Statistics at the University of Florida. He received his Bachelor's degree from the University of Rochester in 1968 and his doctorate from the University of Wisconsin in 1972. Dr. Agresti is a Fellow of the American Statistical Association, and he received an honorary doctorate degree from De Montfort University, U.K., in 1999. He is the author of widely used textbooks, including *Categorical Data Analysis*, *An Introduction to Categorical Data Analysis*, and *Statistical Methods for the Social Sciences*. *The Chicago Chapter of the American Statistical Association* recently named Dr. Agresti its *Statistician of the Year* for 2003. The following conversation took place via email during September and October 2003.

**JD: When and how did you first become aware of statistics as a discipline? When did you decide to study statistics?**

AA: Like most students in my generation who went to graduate school in statistics, I was a math major as an undergraduate. But I had no idea what I would do for a career with a math background. My junior year at the University of Rochester, I took a probability course that used Parzen's classic text. I enjoyed it, and that prompted me to take the follow-up mathematical statistics course. Although that course did not give me much of a sense of what statisticians actually do, it did show me that studying statistics would be a way to apply math skills to problems that had some connection with the real world.

**JD: Did you begin graduate school with the intention of majoring in statistics? How did you select a graduate school?**

AA: Yes, I did apply only to statistics programs. I didn't know much about where the top programs were, but I knew that Wisconsin had an excellent reputation as a university. Also, Madison sounded appealing as a city, and I was attracted by the active antiwar reputation it had. (This was in 1968, during the height of the Vietnam War.) So, my decisions were guided only partly by academic drives. It did help that Prof. G.S. Mudholkar at Rochester told me that Wisconsin had a strong statistics program and that George Box at UW was a nice person as well as an outstanding statistician. But I was in no way yet "committed to statistics," and I was equally concerned in 1968 with American policy in Vietnam and with my own draft status. Besides a fellowship, Wisconsin offered me a teaching assistantship that I



Alan Agresti



Jackie Dietz

could use in appealing my draft status. (At the time, teachers were exempt from the draft. A couple of years later, I was "saved" by the draft lottery that was famous for being unfair; those born early in the year, like myself, were more likely to get high numbers.)

**JD: I was an undergraduate at Oberlin College at that time. I watched that draft lottery on television in the dormitory lounge, surrounded by male students of draft age who were just learning their lottery numbers. I often talk about the draft lottery in my statistics classes, both because the data are so interesting to explore and because it was such a memorable event of my college years. Did you become committed to a career in statistics during your graduate school years?**

AA: Yes, I did, by the end of my four years. There were certainly many times in the first couple of years that I questioned it, such as whenever I attended a seminar and understood very little or was convinced that the field of statistics attracted more than its share of nerds.

**JD: How and when did you decide on an academic career? Have you ever worked as a statistician outside academia?**

AA: I made the decision after I got my Master's degree and decided to go on for a Ph.D. At first, I was nervous about the research pressures that come with a "publish or perish" job in academia. Yet, I enjoyed my experience as a teaching assistant, and I really liked the freedom that comes with academia. I've always had a serious travel bug, so especially important to me was the opportunity (with a nine-month contract) to take more time off in the summers than is possible in the U.S. with most jobs outside academia. My career has been restricted to academia, except for a summer job with the Census Bureau and many visits (one to three days) to various companies to present short courses.

**JD: How do you feel now about that pressure to publish? You've obviously had a very successful**

**research career; have you ever found the research pressure in academia to be excessive or oppressive?**

AA: When I left school, the job I took in the Statistics Department at Florida put strong emphasis on good teaching. Research demands were there, but relatively modest. My position was created to develop statistics courses for students in the social sciences. From working with students and their advisers, I soon realized the social sciences had lots of categorical data. I got interested and changed my research focus completely from the area (branching processes) on which I wrote my Ph.D. thesis to categorical data analysis. Such a change would have been difficult in a university that demanded greater research productivity than Florida did in those days. So, in answer to your question, I was lucky to start my career and develop my research skills in an environment in which the pressure was not excessive.

**JD: Do you think that junior faculty members starting out today are under more pressure to publish and to get grants than we were 25 or 30 years ago?**

AA: They certainly are. But I think there's nothing special about academia in this respect. Job demands and uncertainties are higher throughout the workplace, particularly in the U.S. There is less pressure in academia in other countries, although in many it seems to be increasing and becoming more like the U.S.

**JD: I find it intriguing that your interactions with students and faculty in the social sciences led to a shift in your research area. Can you elaborate about how you learned a new area and how you were able to identify interesting research problems in that area?**

AA: Well, I soon found out that Leo Goodman was "God" to quantitative social scientists. Goodman, a statistician then at the University of Chicago, has been the most prolific researcher over the past 50 years in categorical data analysis. For each article he wrote in a statistics journal, he wrote one on the same topic but at a more applied level for a social science journal. This is not a bad model for statisticians to emulate! So, I spent a lot of time reading his articles. It took a few years, but gradually I got a sense of the state-of-the-art research by him and by others who worked in this area, such as Gary Koch and Steve Fienberg. I also learned a lot from the 1975 classic text *Discrete Multivariate Analysis*, by Yvonne Bishop, Steve Fienberg, and Paul Holland, which gives an elegant presentation of log-linear models for contingency tables. My own research was often motivated by questions I got while interacting with colleagues at Florida, such as "I know about Fisher's exact test for 2x2 tables. What can be done with this

table that has more than two categories and which are ordered?" Overall, my research was helped by working in a still relatively undeveloped area and mainly on real rather than artificial problems.

**JD: I'm interested in how your teaching, consulting, and research have motivated and influenced each other. You've said that working with students and faculty from the social sciences piqued your interest in categorical data. Do you continue to get research ideas from statistical consulting?**

AA: Yes, occasionally. In fact, I'm currently working on a paper in response to a question I've received twice recently while visiting pharmaceutical companies. In magazine advertisements for new drugs, you'll often see summary tables that compare the relative frequency of each of several adverse side effects for the drug and for a placebo, based on results from placebo-controlled clinical trials. How can one conduct a global test of equality of the vector of population proportions for the drug and the vector of population proportions for the placebo? For multivariate normal responses, Hotelling's  $T^2$  tests equality of two vectors of means. For multivariate binary data, there are many possible answers to the question, but none is entirely satisfactory. Methods can be computationally intensive or asymptotic inference can be inadequate when each vector has a large number of elements, because of data sparseness.

**JD: Do you incorporate the results of your research into your teaching and consulting?**

AA: I try to. Even in teaching elementary courses, I try to give students some historical perspective and explain to them that statistics, like any field, is continually evolving. I always mention a few important modern advances, such as the bootstrap, but once or twice I'll try to briefly say a bit about what I do or have done in statistics research. Students are surprised that you can do research in statistics. They imagine that we teach a tool bag of methods that have been in existence for hundreds of years. Sometimes teaching itself can even inspire research. Six years ago I started some work on binomial confidence intervals after wondering about the sample size guidelines for the simple confidence interval for a proportion that we teach in every basic statistics course.

**JD: I know you've shown that adding two successes and two failures to the sample before calculating the usual Wald confidence interval for a proportion yields coverage probabilities that are closer to the nominal confidence levels than those of the unadjusted Wald interval. Do you now teach that method in your introductory courses? Will you include that method in the introductory text you're**

**currently writing? This seems like a perfect example of a recent research result that is simple enough to teach to beginning students.**

AA: I do mention this briefly, partly to show students that although the **estimate  $\pm$  2 standard errors** formula is versatile, it sometimes breaks down. Students respond better to an example than a derivation. In collecting data for one course, I asked the 25 students if they were vegetarian. No one was, and when they used the ordinary Wald confidence interval and got (0, 0) for the population proportion, they realized it was nonsensical.

It's been gratifying to me to see that some introductory texts now recommend this simple "add 2 successes and 2 failures" interval that Brent Coull and I proposed (*The American Statistician*, 1998). These include the texts by Moore and McCabe, McClave and Sincich, and Samuels and Witmer. In the past, texts at this level have not told students how to construct a confidence interval for a proportion when the sample size is small or when relatively few observations occur in one category. I think that our interval gives a simple solution for courses in which discussion of more complex methods (e.g., score interval or likelihood-ratio intervals) would be beyond the scope. And yes, thanks for asking, this method will appear in the upcoming introductory book by Agresti and Franklin.

**JD: Agresti and Franklin will be the fifth textbook you've written, following three books on categorical data analysis and Statistical Methods for the Social Sciences. Writing five books seems like a daunting task to me—what factors have motivated you to write textbooks?**

AA: Well, I think it's natural for any teacher to be unhappy with certain aspects of any text they use and to feel they can do it better. My first book, for the social sciences, was motivated by seeing that most texts for that audience had serious deficiencies. Perhaps this is because they were written by social scientists rather than statisticians. For instance, I remember one that presented a null hypothesis for comparing means as  $H_0: x_1 = x_2$ . These days, there are excellent books in introductory statistics, such as those by David Moore. But there's always room for a different slant, such as increasing the use of simulations and activities. Of course, whether authors can pull off well what they envision is never guaranteed!

**JD: What have you enjoyed most (and least) about the process of writing a textbook?**

AA: It's a nice complement to other professorial work. For instance, with research you can have bad periods in which you don't seem to be making progress or the problem you're working on is not that exciting. In writing a book, with every hour of work you can feel that you're making some progress. And, of course, it's

very satisfying when the book comes out, and then later when a royalty check arrives. I've also felt that writing a book helps me to broaden my knowledge and to organize my thoughts about a subject area. Lastly, if a book is successful, you get some nice feedback and more recognition than from your research and teaching – probably more than you really deserve. The hardest part for me is that with such a large project, you have to fight to keep it from taking over your life. You can be watching a movie at night, and your mind wanders to that section you wrote today that really could be improved.

**JD: I think one of the toughest challenges of writing a book would be finding datasets for examples and problems. Do you consider it essential to use real data in your books? If so, how do you find appropriate datasets?**

AA: I do prefer to use real data, when possible. This is much, much easier than it used to be, because of the internet. Large databases are increasingly available. An example is the General Social Survey, a survey conducted of Americans every couple of years (see [www.icpsr.umich.edu/GSS](http://www.icpsr.umich.edu/GSS)). You can go to their web site and with a few clicks download information on diverse topics such as belief in heaven and hell, time spent watching TV every day, the proportion of people who consider themselves happily married, opinions about controversial issues such as abortion and affirmative action, and so forth. Also, search engines make it increasingly easy to find summary results on nearly anything that interests you. I was looking for a golf example yesterday, and a one-minute search with Google led me to a multiple regression equation that predicted total scores in a Masters golf tournament based on predictors such as the number of putts and the number of greens reached in regulation. Of course, with such searches it's usually a lot easier to find summaries of data than the raw data themselves.

**JD: You mentioned earlier that you like to give your students some historical perspective and an awareness of modern advances in statistics. What do you think the most important recent developments in the field have been? What do you think will be the most exciting and productive areas of research in statistics during the next few years?**

AA: During the past quarter century, biomedical applications have provided a variety of interesting topics for research, such as models for censored survival data, designs for clinical trials, and methods to handle missing data. I imagine that this application area will continue to generate research problems, especially with the increasing importance of genetics. A second area that has continually increased in importance since the introduction of the bootstrap is computationally intensive statistics. When I was a graduate student, many seminars

began “Let  $X_1, X_2, \dots, X_n$  be an i.i.d. sequence of random variables” and then never seriously considered any application or actual data. I think that the field will continue to move away from that in the direction of considering models (especially hierarchical ones) motivated by nonstandard data structures and large multivariate datasets. Problem areas where such datasets are common include data mining and credit scoring in business, and bioinformatics—for example, the analysis of DNA sequences. Now, I imagine that attention will also be devoted to databases and analyses relevant to “homeland security.” So, as usual, I would guess that the most exciting and productive areas of future research will be ones generated by application areas that are themselves very active.

**JD: What do you see as the greatest challenges facing the profession of statistics in the coming years?**

AA: There’s been lots of insightful discussion of the many challenges in the annual addresses by the presidents of the American Statistical Association, as published every year in *JASA*. I’ll just mention one. The demand for statisticians continues to increase, but the number of students from this country who are going on to obtain Master’s and Ph.D. degrees is not. Every statistics department in the U.S. has trouble finding American students, and I think the same applies to Canadian students and British students in those countries. Partly this reflects the severe drop over the years in undergraduate math majors. It would help immensely if there were more undergraduate programs in statistics and if the existing ones were modernized to make them more attractive. We have a wonderful opportunity currently, as more and more high school students are taking statistics and passing the Advanced Placement exam. So, we have more students coming to colleges and universities aware that the field of statistics exists. Otherwise, we don’t have contact with most students until they are juniors or seniors and have to take statistics as a required course for their major. It’s then too late.

**JD: Will user-friendly statistical software reduce the demand for our expertise?**

AA: I really don’t think so. About 20 years ago some predicted that expert systems would be developed that would make us redundant, but it just hasn’t happened. This is no surprise to anyone who does any consulting. It’s highly interactive. You need to ask a lot of questions and find out what the main issues are that need to be addressed before you can think about recommending a particular method. And then, there’s continually a greater and greater variety of possible methods, and some depth of understanding is needed to give good advice about what’s appropriate.

**JD: Will academic departments of statistics become less necessary as researchers in other disciplines acquire more quantitative skills?**

AA: Again, I don’t think so. As more people in other areas realize the need for statistics and develop some expertise in it, I think this makes them appreciate greater contact with professional statisticians having similar interests and possibly deeper knowledge than they have the time to develop. But this brings up another challenge for us in the future—doing a good job communicating with nonstatisticians, something for which our field does not have an especially good reputation. It’s especially important that we maintain and teach well relevant courses for graduate students in other areas, starting with basic statistical methods and including a few courses on advanced topics.

**JD: You mentioned the importance of undergraduate statistics programs in meeting the demand for statisticians. Do you see those programs primarily as a source of graduate students in statistics, or do you think there are attractive employment opportunities for statisticians at the Bachelor’s level?**

AA: Undergraduate programs can be an important source of graduate students, but I don’t see that as their primary reason for existence any more so than for undergraduate sociology or marketing or any other program. If we believe that statistics is an important subject in its own right, then it should be worthy of its own program. Of course, in a small college or university, this program might be a concentration within another department, such as math. I don’t know how well-defined the employment market is for Bachelor’s degrees in statistics, but given the heavy use of statistics these days, there should be opportunities. The jobs may not always be advertised as “statistician,” but that’s what many jobs end up being that advertise for computing or other quantitative skills. And let’s not forget the important need for more statisticians to teach at the high school level.

**JD: Do you have any advice for students considering an undergraduate statistics major?**

AA: It’s good to have a mix of theory and application, such as two-term sequences in mathematical statistics and in linear models. If you plan on graduate school in statistics, try to take advanced calculus and matrix algebra. If you plan to go directly to a job, you might include more computing and some specialized courses in topics such as multivariate analysis, biostatistics, categorical data analysis, and statistical computing.

**JD: What advice would you give students who are planning to go to graduate school in statistics? Any tips for selecting a graduate school? Are there particular elective courses you think every graduate student should take? Do you have any suggestions for choosing an adviser?**

AA: I don't think I have any special insight here. Try to get into one of the top programs for which you are prepared. It's usually good to take courses from, and have as an adviser, a faculty member who is distinguished in the field. That being said, often young faculty members who are not yet well known have more time to devote to their students. (This was my experience, as I was lucky to have Steve Stigler as my adviser at Wisconsin, shortly after he received his Ph.D.) I'm reluctant to pick elective courses that every student should take. Lots of what was "hot" when I was a student seems not so fundamental now, and what is "hot" now was not around then. I think it's mainly important to get a strong background in the fundamentals of statistical theory and modeling. The modeling should not be solely "normal-based" but include the variety of types of data seen in practice, such as binary data, count data, and censored data.

**JD: I didn't realize that Steve Stigler was your adviser! Were you his first student? Are there other people or events that have been influential in your career?**

AA: I was his first student to finish, in 1972. Soon after, he had other excellent students who eclipsed me, such as L.J. Wei and Rob Kass. I can't give enough credit for how helpful he was to me, from the generosity with his time to helping me get the job at Florida.

Since leaving graduate school, the most influential work-related event was having my first sabbatical year at Imperial College, London, during its heyday in 1981-1982. It had an outstanding young faculty, lots of interesting visitors because of the presence of D.R. Cox, and London had considerable activity at other universities as well. It was a rewarding year, one where I learned a lot about new topics such as generalized linear models and GLIM. This led to London being my second home, these days less for statistics than for everything else it has to offer. For my next career, I'd be quite happy to be a theatre, music, and restaurant critic in London!

The most influential non-work-related event was meeting

my wife, Jacki Levine, in 1986. She helps me keep things in perspective, and she has her own career (first newspaper journalist, now managing editor) that keeps me reminded that there are other interesting careers and more to life besides statistics!

## Books Mentioned in the Interview

- Agresti, A. (1996), *An Introduction to Categorical Data Analysis*, New York: Wiley.
- Agresti, A. (2002), *Categorical Data Analysis (2nd ed.)*, New York: Wiley.
- Agresti, A., and Finlay, B. (1997), *Statistical Methods for the Social Sciences (3rd ed.)*, Upper Saddle River, NJ: Prentice-Hall.
- Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975), *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, MA: MIT Press.
- McClave, J. T., and Sincich, T. (2003), *Statistics (9th ed.)*, Upper Saddle River, NJ: Prentice-Hall.
- Moore, D. S., and McCabe, G. P. (2003), *Introduction to the Practice of Statistics (4th ed.)*, New York: W. H. Freeman.
- Parzen, E. (1960), *Modern Probability Theory and Its Applications*, New York: Wiley.
- Samuels, M. L., and Witmer, J. A. (2003), *Statistics for the Life Sciences (3rd ed.)*, Upper Saddle River, NJ: Prentice-Hall. ■

### M.A. in Quantitative Methods in the Social Sciences (QMSS)



**Columbia University**



Columbia University offers an interdisciplinary M.A. degree program that trains students in how to apply quantitative methods to a variety of issues in non-profit organizations, government, business, and social research. The program combines rigorous training in statistical techniques with an examination of how these methods are applied to a diverse set of problems in the social world.

The program is structured for both full-time and part-time students. Past graduates have found positions in non-profits, market research, public health, finance, and government. Individuals working in these areas would also benefit from the specialized training in quantitative methods this degree provides.

Applications to begin study in fall 2004 are due May 1.

qmss@columbia.edu

www.qmss.columbia.edu

212-854-8039



# Statistics Projects with a Future

Ever thought of presenting the results from your statistics project at an academic conference? The first author and three of his students presented papers at the 2000 International Applied Business Research Conference in Puerto Vallarta, Mexico. You, too, can do this.

We encourage you to entertain the thought of presenting the results of your research at a conference. Ideally, this can lead the statistics project to blossom into one of the most memorable learning experiences of your college career. While the focus of this article is on some of the mechanics of, suggestions for, and benefits of conference presentations of your projects, remember that the focus of research should be on answering questions truly of interest to the researcher. Conference presentations should not be an end in and of themselves, but a step along the path in the quest for knowledge.

## What to Present

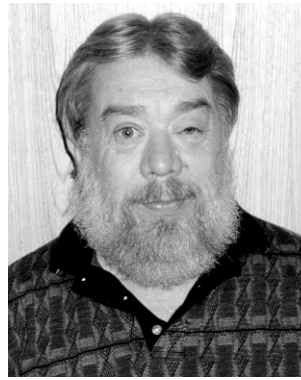
You may have already completed a statistics project that is worthy of presentation at a conference. Does your project provide any new information or insight useful to the academic community, or does it add to a growing quantity of research on a popular topic? The information need not be overwhelming, but it should be original. Different disciplines vary greatly in the topics they are interested in, so finding the most appropriate one may take some looking. Reviewers' opinions also vary greatly, so do not give up with one rejection. Try to interpret your results in the way most useful to your audience.

If you don't have an appropriate project, you might look for one that could provide information worthy of being published. Questionnaire surveys are a popular method for business and social science

---

*George Kesling was a professor of business administration at Central Washington University. He earned his Ph.D. from the University of Oregon, and he taught statistics and computing there and at the University of Hawaii, the University of Washington, and the University of Wisconsin.*

*Wayne Fairburn is a professor of business administration at Central Washington University where he has taught finance and statistics for many years. He received his Ph.D. from Michigan State University.*



**George Kesling**



**Wayne Fairburn**

students. Choose a topic that is of interest to you, and of common interest beyond your school. The academic community is interested in topics such as factors associated with student grades, cheating on tests, and timely topics such as factors associated with internet purchases. More widely read media are interested in a wide variety of topics if someone has some new measurements. Good measurements and interpretation are key factors in the academic value of statistical projects, and the topics are limitless.

## Where to Present

If you feel your project may be worthy of presentation, try to find conferences that include your project topic and are at acceptable times in desirable locations. Search the internet. Start with [www.decisionsciences.org](http://www.decisionsciences.org), [www.amstat.org](http://www.amstat.org), [www.allconferences.net](http://www.allconferences.net), [www.papersinvited.com](http://www.papersinvited.com), [www.appliedsoc.org](http://www.appliedsoc.org), [www.apa.org](http://www.apa.org), [www.maa.org](http://www.maa.org), and [www.nsns.com/MouseTracks/](http://www.nsns.com/MouseTracks/), or other sites you or your teachers may have seen. Look at the "calls for papers or participation" and note the deadlines. The conference must be several months in the future to allow you time to revise your class paper for a new audience and meet the submission deadlines, which usually start about three months in advance. Choose a conference that appeals to you. Remember, there are local, district, national, and international conferences, and each has advantages and disadvantages. If the conference publishes its proceedings, then there can be an ISSN number in the Library of Congress of a document containing your paper. This is the best documentation, so you may want to choose a conference that publishes its proceedings. You risk nothing but your time if it doesn't work out.

Revise the class report for your chosen conference. Look at past papers presented at the conference, if you can. Use the same format and try to relate your paper to the topic or theme of the conference. Plan to submit your paper to the student competition section if there

is one. Otherwise, plan to submit it to the general competition in the section most appropriate for your topic. Since most submissions are usually from faculty, professionals, and a few graduate students, most conference participants like getting a few student perspectives. Most judges (reviewers) also like hearing from articulate students. Use your natural strengths. One student of ours was only average in statistics, but exceptional in presentations. She presented a slide show of the students doing their projects and the audience really liked her presentation. Another student got a trip to Hawaii mostly paid for with a well-articulated description of the frustrations he felt doing his statistics. Usually, however, our students' papers have been based on good interpretations of measurements collected on topics such as factors associated with student grades, internet purchases, and computer usage and ownership.

Submit your paper. This is the critical step that sets you above 90% of the academic community. Most conferences consider drafts of papers that are expected to be completed by the time of the conference. Usually, the more complete the paper, the better the chance of acceptance. Most importantly, submit what you have by the deadline. If your paper gets rejected, use the reviewer's comments to revise it and submit it to another conference. If it gets accepted, you have already gained some recognition.



## Getting There

Once your paper has been accepted, start investigating funding possibilities to help pay your expenses, which usually include a conference fee and travel expenses. Use your acceptance letter to get funds from a variety of sources. Plan to pay part of the expenses yourself. Look for bargains. Make a realistic cost estimate. Travel expenses usually include transportation, lodging, food, and other expenses. By

getting some contributions from a variety of sources, you may get most of your costs paid.

Your school can gain recognition when you present a paper at an academic conference; most accrediting agencies emphasize evaluation of a school by its "outputs" which are, of course, its students' performance, so the school may be willing to pay part of your travel expenses. Seek the help of your teachers and department advisers. Does your project qualify for funds from any minority or other special interest group?

Make arrangements in advance. Try to get there at least one day early. Find the room where you will give your presentation in advance and stand where you will be presenting, so when you actually make the presentation you will have been there before. Most of the work will be done before you get to the conference. Try to relax and enjoy the experience. It is likely to get you "all fired up."

## Being There

Show up and deliver the goods. Don't read your paper; give your audience a copy and talk about what it says. A few transparencies or slides will take the focus off of you and onto the project. Expect a dozen or more people in the audience. If someone asks you a question and you don't know the answer, say: "That is a good question; I will try to find the answer and get back to you." Don't be afraid to admit that you don't know the answer.

Have fun after you're done. This is the time to celebrate your achievement!

Document your achievement. Be sure to put it on your resume. Conference presentations are something most teachers are expected to do, and presenting as a student is particularly impressive. Usually, teachers, administrators, or other friends will help you and your school get the recognition you deserve. If no one else does it, then do it yourself.

When asked, "What is your most memorable experience in college?" we hope some of you will start your answer with, "Well, we did this statistics project..."

## Web Resources

American Psychological Association: [www.apa.org](http://www.apa.org)

American Statistical Association: [www.amstat.org](http://www.amstat.org)

Decision Sciences Institute: [www.decisionsciences.org](http://www.decisionsciences.org)

Lists of Conferences: [www.allconferences.net](http://www.allconferences.net),  
[www.papersinvited.com](http://www.papersinvited.com)

Marketing: [www.nsns.com/MouseTracks/](http://www.nsns.com/MouseTracks/)

Mathematical Association of America: [www.maa.org](http://www.maa.org)

Society for Applied Sociology: [www.appliedsoc.org](http://www.appliedsoc.org) ■

# Play ASA Stat Bowl in Toronto!

The ASA Stat Bowl will be back at the Joint Statistical Meetings (JSM) in Toronto in 2004. This event was a big success at JSM 2003 in San Francisco, and we're hoping the same will be true this time around.



Just like last year, each student who plays in Stat Bowl will receive up to \$500 in travel reimbursement, regardless of his or her performance in the Bowl. Don't let the cost of travel keep you home this August—the Stat Bowl can help get you across the border! Don't worry if you are the only student in your program interested in playing. Stat Bowl is an individual competition. Team points are kept by school and a team championship is awarded, but having a team is not a requirement.

The contest in San Francisco was a lot of fun. Wes Eddings from Johns Hopkins University was the individual champion (see pg. 29). He exhibited extraordinary knowledge of many different aspects of statistics. David Hitchcock from the University of Florida also impressed the audience and was awarded the runner-up trophy. Propelled by David's performance, the team from Florida won the team championship, defending the crown Florida won in 1999, when the College Bowl was last held at JSM in Baltimore.

Students will be accepted into the 2004 tournament on a first-come, first-in basis. Notification of a willingness to participate will serve as entry. Inquires about the Bowl or requests to be registered as a contestant can be sent to me at [mpayton@okstate.edu](mailto:mpayton@okstate.edu). A maximum of 16 players will be allowed in the contest. In the event that the field of contestants fills to capacity, each university will be restricted to two players to assure diversity. A waiting list will be established to fill unexpected vacancies at game time.

The contest will be held on the Tuesday of JSM in two sessions. Session 1 will consist of four games, each with four contestants. The winners of these four games, plus two at-large contestants, will advance to Session 2. Players who score the most points in Session 1 without winning their game will be the at-large contestants. Session 2 will consist of two games, each with three players. Those two winners will then meet head-to-head in a championship game.

---

Mark Payton is a professor in Oklahoma State University's department of statistics.



**Mark Payton**

Many different rounds make up a Stat Bowl contest. There's the Toss Up/Bonus round (which consists of bonus questions awarded to players that buzz in and answer a correct toss up question), Lightning Round (quick questions that have no bonus questions attached), Category Round (players pick a category and receive questions and bonus questions), and the Final Round (players write their answer to the last question of the game). The questions focus on the ASA organization and on statistical history and methodology. Some sample questions appear in the sidebar.

Any student interested in playing should contact me ASAP before the player positions are filled. We encourage all players to register before July 1, 2004, but players will be added up to game time if space is available. Hope to see you in Toronto! ■

## Sample questions:

1) TOSS UP: A discrete random variable  $Y$  has the following probability function:  $p(y) = cy, y = 1, 2, 3, 4$ . What is  $c$ ?

2) BONUS: What is the expected value of  $Y$ ?

3) CATEGORY ROUND: This man was a disciple of Chebyshev. His outstanding contribution to probability theory was the introduction of a concept that served as a model for the study of dependent random variables, which is at the heart of a vast amount of research in the theory of stochastic processes. Name this man famous for his "chains."

4) LIGHTNING ROUND: Who is the 2003 editor of *CHANCE* magazine?

5) FINAL ROUND: Although Susie had a better batting average than Freddie in each of the previous five seasons, it is possible for Freddie to have a better overall batting average for the five seasons. What is this paradox called?

*answers on page 29*

# When Direction Vanishes: Walking Straight or in Circles

Every year, there are news reports about families or individuals who get lost in the desert, mountains, or woods (e.g. *Arizona Daily Star*, 2002 and *Las Vegas Review Journal*, 1998). These people wander aimlessly, only to be found close to a trail or to a well-traveled road. Frequently, these same people report they thought they had wandered in circles for hours before help arrived. Several famous children's stories, including the tale about Winnie-the-Pooh and Piglet, who get lost in the Hundred Acre Woods, reference characters that wander in circles when all direction vanishes (London, 1999 and Paulsen, 1985). I designed a study to investigate if, when an individual does not have a defined path or compass and cannot see his/her final destination, she or he will deviate from the direct path and instinctively curve to the right or to the left rather than walk in a straight line. Individuals in this study were blindfolded; each had no knowledge of where she or he was going, to represent this scenario.

The purpose of this study was to demonstrate whether, despite an individual's best attempt to walk in a straight line, human physiology causes the person to walk in a circular or elliptical pattern. The study also investigated whether or not a person's dominant hand dictates the direction of the curve, if the height of the person correlates to the angular acceleration of the curve, and if there is any difference between men and women under the same conditions.

## Study Design

A total of 32 people, 16 men and 16 women, were recruited from Sabino High School in Tucson, Arizona. Students were selected at random and asked to participate in this study. The subjects were selected independently and the selection of men did not influence the selection of women.

Each participant was asked to sign a consent form and to complete a brief information sheet on which

---

*Andrea Axtell (jaaxtell@aol.com) is an honor student at James Bowie High School in Austin, Texas. She credits her love of math and science to her mother and grandmother, who continues to remind her that "math is for girls, too." Andrea is also a world-ranked competitive swimmer and member of the Longhorn Aquatics National Championship Team.*

they were asked to record their gender and whether they were right- or left-handed. The person's height was measured using the Athletic Department's medical scale and recorded (in inches) on the individual's test form.

The walk tests were conducted on the football field adjacent to the high school. The school and all of its facilities are situated entirely within a quiet residential neighborhood. Access to the school is limited to one driveway, and the football field is located at the back of the property, more than a mile from a major street or controlled intersection. The tests were conducted during the day. Vehicular traffic in the school's parking lots was almost nonexistent.

Each person was tested individually while the other participants remained in the gym or distant practice fields. Subjects were not permitted to watch the other subjects' walk tests. The walk test started on the center hash mark of the goal line closest to the school. Each subject was asked to square his/her body with the goal post at the other end of the field.

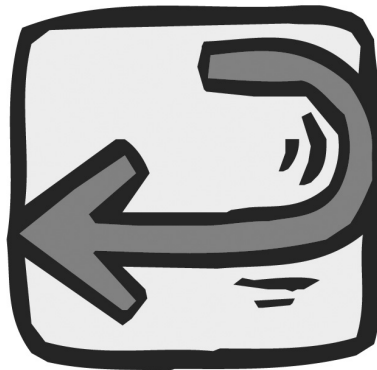
Control subjects were asked to walk forward in a straight line toward the center hash mark on the opposite goal line or until they were told to stop. Experimental test subjects were blindfolded with a dark colored polar fleece hat and scarf. After the blindfold was placed over the person's head and face, he or she was asked to walk forward in a straight line until told to stop.

As the subject crossed each five-yard line, I placed a marker (purple streamer tied to a coat hanger) in the ground. The field was marked every five yards until the person reached the opposite goal line or until the person reached one of the football field's defined boundaries or sidelines. When the subject reached one of these points, he or she was told to stop, and the blindfold was removed.

At each five yard point, the distance between the center of the field, as designated by the center hash mark and the respective streamer, was measured and recorded. Each person's specific walkpath was plotted and graphed.



**Andrea L. Axtell**



Handed	Height	Yard line	Gender	Handed	Height	Yard line	Gender
right	66	52	male	right	63	40	female
right	74	95	male	right	64	42	female
right	69	60	male	right	70	65	female
left	70	68	male	right	65	48	female
right	66	56	male	right	68	59	female
left	64	49	male	right	71	71	female
right	64	37	male	right	65	52	female
right	70	73	male	left	63	38	female
right	71	70	male	right	62	37	female
left	68	61	male	right	66	55	female
right	73	80	male	left	65	50	female
left	70	70	male	right	64	42	female
right	70	74	male	left	72	81	female
right	68	60	male	right	67	56	female
right	70	75	male	right	61	35	female

## Results

Data collected on each subject appear in Table 1. Height is reported in inches, and the third column refers to the yard line at which the subject crossed the sideline (subjects with a higher value walked further before deviating off course).

Two control subjects, who were not blindfolded, did not deviate from the centerline and walked in a straight line to the designated center hash mark at the end of the football field. None of the experimental test subjects ( $n = 30$ ) walked in a straight line to the designated center hash mark at the end of the football field. Fully 100% deviated from the centerline, angled toward one side of the field or the other, and eventually crossed the sideline before reaching the designated destination. A few subjects deviated from side to side but then angled toward the sideline. Other subjects deviated from the centerline almost immediately and walked in a nearly perfect arc toward the sideline. A smaller number walked straight for a long time before they too angled toward the sideline.

### Point Where Subject Crossed Sideline

The point where experimental test subjects crossed the sideline ranged from 35 yards to 95 yards. The mean point where subjects crossed the sideline was 58.4 yards. The median point where subjects crossed the sideline was 57.5 yards. The distribution of these yard lines is shown in Figure 1.

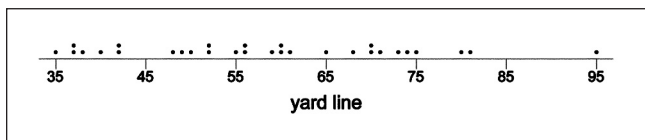


Figure 1: Distribution of Yard Lines Where Subjects Crossed Sideline

In an effort to determine if men crossed the sideline at a distance further from the goal line than women, summary data were analyzed to compare the points where male test subjects crossed the sideline

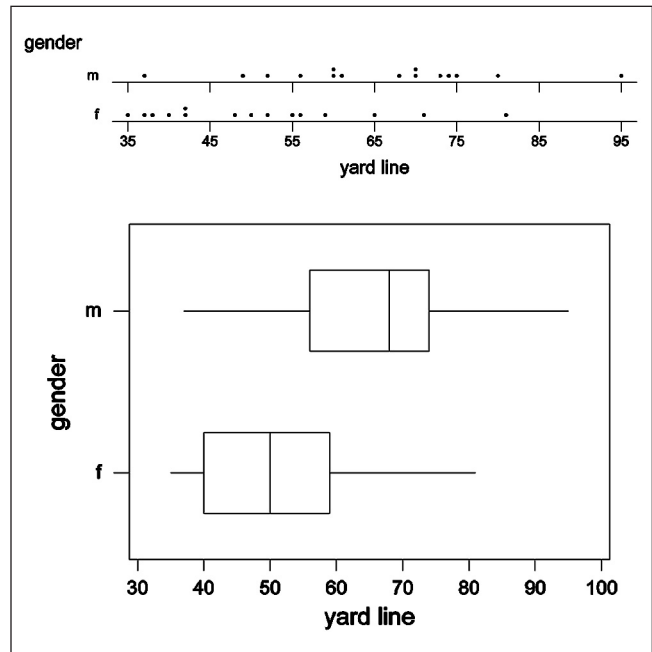


Figure 2: Distributions of Yard Lines Where Subjects Crossed by Gender

with the points where female test subjects crossed the sideline. These distributions are shown in dotplots and boxplots in Figure 2. The majority of males (80%) crossed the sideline between 51 and 80 yards; the majority of females (80%) crossed the sideline between 31 and 60 yards. The mean for males was 65.3 yards, with a median of 68 yards. The mean for females was 51.4 yards, with a median of 50 yards.

Using the two-sample  $t$ -test for comparing the means of two independent populations, the test statistic is  $t = 2.77$  and the  $p$ -value is .010. This suggests there is sufficient evidence to conclude that males tend to cross the sideline at a distance further from the goal line than the point where females cross the sideline.

### Right-Handed Subjects vs. Left-Handed Subjects

Twenty-two of the twenty-three (96%) right-handed experimental test subjects crossed the right sideline, while one right-handed subject (4%) crossed the left sideline. All seven left-handed subjects (100%) crossed the left sideline. Based on a chi-squared distribution with  $df = 1$ , the computed  $X^2$  value (25.1) produces a  $p$ -value of less than .001. The use of the chi-square test is suspect here due to the small expected counts, but the data provide strong evidence that right-handed individuals tend to deviate to the right sideline and left-handed individuals tend to deviate to the left sideline.

### Height

Measuring a person's stride was outside the confines of this experiment, so I assumed the person's

height was directly proportional to the person's stride and those who were taller would have a longer stride than those who were shorter. I tried to determine if there was any difference between the point where taller subjects crossed the sideline and the point where shorter subjects crossed the sideline.

Figure 3 displays a scatterplot to provide a visual impression of how strongly the height of the subject and the point where the subject crossed the sideline are related. The correlation coefficient was calculated ( $r = .975$ ). The data suggest a substantial positive linear relationship between the point where the subject

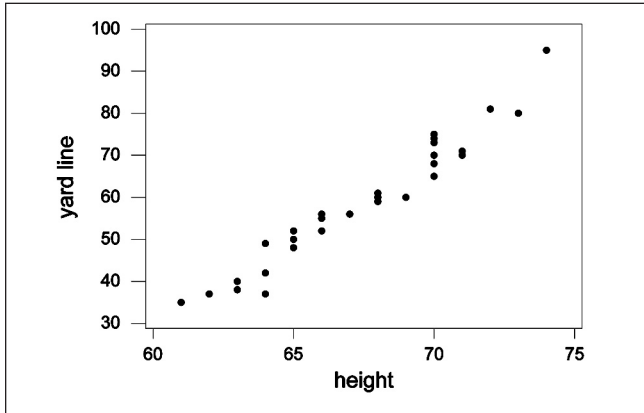


Figure 3: Relationship Between Yard Line Where Subjects Crossed and Height

crossed the sideline and the subject's height. The p-value for testing whether the correlation coefficient differs significantly from zero is less than .001.

## Height and Gender

As noted above, the results demonstrate a statistical difference between the point where males cross the sideline and the point where females cross the sideline. The data also demonstrate a significant relationship between the subjects' height and the point where they cross the sideline.

Combining this information into one analysis, Figure 4 displays a scatterplot of yard line vs. height, with different symbols for men and women subjects. Figure 5 shows the same plot with two least-squares

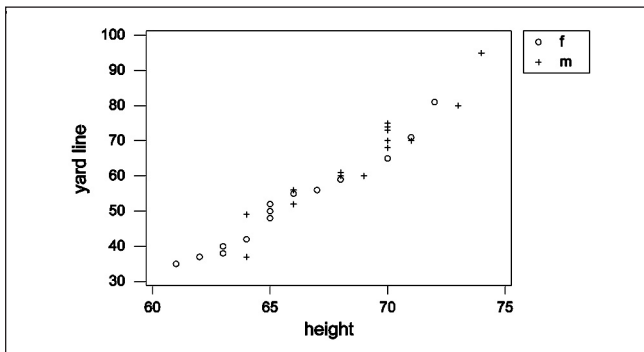


Figure 4: Relationship Between Yard Line Where Subjects Crossed and Height, by Gender

lines drawn in: one for predicting a male's yard line from his height and one for predicting a female's yard line from her height. The equations of these lines are:

- Males: predicted yard line =  $-251 + 4.60 \text{ height}$  ( $r^2 = 91.0\%$ )
- Females: predicted yard line =  $-212 + 4.01 \text{ height}$  ( $r^2 = 96.7\%$ )

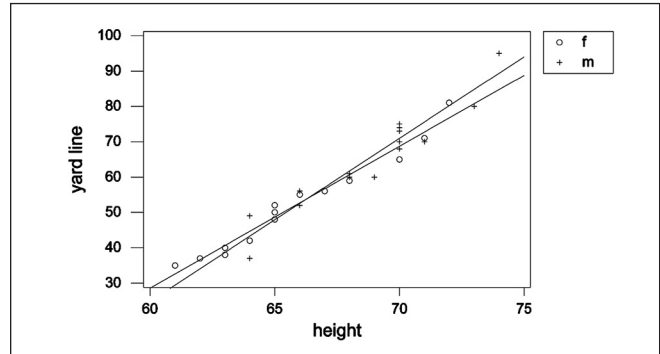


Figure 5: Least Squares Lines for Predicting Yard Line from Height by Gender

Since these lines are so similar, it reveals the gender effect is not significant after one adjusts for height. (A multiple regression analysis including both explanatory variables produces a p-value on height of .000 and a p-value on gender of .717.) In other words, males tended to walk further before deviating off-course than females, but that difference can be explained by males tending to be taller than females.

## Summary and Conclusion

The purpose of this study was to determine whether, when an individual does not have a defined path or a compass and cannot see his or her final destination, he or she will deviate from the direct path and angle to the right or to the left rather than walk in a straight line. It was designed to determine which direction right-handed subjects would turn and which direction left-handed subjects would turn. The study was designed to determine if the person's height affected the angular acceleration of the curve and therefore the point at which the subject crossed the sideline, and it examined the differences between men and women under the same conditions.

The data support the hypothesis that if direction vanishes, then people do not walk in a straight line. These findings agree with exercise physiologists' analyses of a runner's stride and support the effect of brain dominance on gross motor movement. When the subject walked without knowledge of where she or he was going, the sideways thrust of the foot caused the subject to angle from the centerline. For the purposes of this study, an assumption was made that all of the subjects were healthy and exhibited normal balance and equilibrium. Only those subjects who were able to see where they were going were able to make the

minor corrections necessary to walk in a straight line to the designated destination.

It was found that people curve in the direction of their dominant side. Right-handed people angled to the right and left-handed people angled to the left. Interestingly, the one subject who said he was right-handed but who angled to the left said he pitched baseball with his left hand and batted left, but was taught to write with his right hand. Since research states that the writing hand is not necessarily the dominant hand, it is possible this subject was actually left-dominant and had only learned to write with his right hand and therefore considered himself "right-handed." The data provided sufficient evidence that the person's dominant side dictated the direction the subject turned.

Since measuring the person's stride was not possible within the confines of this study, the subject's height was substituted. The experiment demonstrated that subjects who were shorter deviated from the centerline more quickly and crossed the sideline at a shorter distance from the goal line than taller subjects.

Experimental errors may have occurred when the subject's height was substituted for the subject's stride in determining the distance at which the subject deviated from the centerline and eventually crossed the sideline. Research tells us stride is affected by leg length, level of physical fitness, and emotional state or purpose. Although the correlation coefficient measures the strength of the linear relationship between two variables, a high correlation does not necessarily imply a cause-and-effect relationship. However, further statistical analysis did demonstrate a significant relationship between height and the point where subjects crossed the sideline. Further research and testing of stride would confirm these findings.

Initial analysis led one to predict that males and females deviated from the centerline and crossed the sideline at different points. However, when gender was paired with the subject's height, it was determined there was no statistical difference between the walkpath of men and women after adjusting for height.

Further mathematical modeling suggested that if the subjects had continued walking they could have completed a full circle or ellipse.

This research has applications for outdoor enthusiasts, search and rescue teams, and possibly military strategists. While one should always carry a compass or a map when traveling in unknown territories, simple strategies could be developed to avoid walking in circles. When trees or rocks block a person's path, people could be trained to alternate going to the right or to the left to avoid further curving to the dominant side. Knowing whether a person is right-handed or left-handed, as well as the person's height, could help define search areas.

Did Pooh and Piglet walk in circles in the Hundred Acre Woods? This study concludes that ... yes, they probably did.

## References

"Family of Four Found in Desert," *Arizona Daily Star*, August 21, 2002.

London, J. (1999), *To Build a Fire*, UC Regents.

Paulsen, G. (1984), *Tracker*, Simon and Schuster.

"Teen Snowboarder Rescued on California Mountain Dies," *Las Vegas Review Journal* and *Las Vegas Sun*, 6B, February 22, 1998. ■



# Customer Service Center

Do you have an address change, membership question, claim, or general inquiry? Please call ASA's new Customer Service Toll Free Direct Line 1-888-231-3473 for one stop shopping for all of your ASA needs. If you prefer, you may email Customer Service at [memdept@amstat.org](mailto:memdept@amstat.org) or fax 1-703-684-2037.

**1-888-231-3473**

**Customer Service • Toll Free • Direct Line**

# Can Major League Baseball Teams Buy Success?

The 2003 World Series pitted the New York Yankees, a franchise famous for its hefty payrolls, against the Florida Marlins, one of the more frugal clubs in recent years. The Yankees have made regular trips to the World Series and are always among the top-spending clubs. The Marlins' only other appearance in the Series was in 1997 when the owner at that time spent huge sums of money to attract a star-filled roster, only to dismantle the team by selling off those high-paid players, producing a dismal 54-108 record the following year. Such examples have helped foster a popular notion that the "rich" teams have a big advantage in putting together a successful ball club, while "small market" teams are doomed to mediocrity. A recent popular book, *Moneyball: The Art of Winning an Unfair Game* (Lewis, 2003), discusses how one team, the Oakland Athletics, has tried to use creative statistical analysis to combat this perceived monetary disadvantage. In the latest collective bargaining agreement, Major League Baseball (MLB) owners instituted a "luxury tax" to penalize teams with extravagantly high payrolls and provide additional funds for the poorer teams. But is there really a relationship between spending and success?

From a statistical viewpoint, the question of interest is whether data support a claim that a team's success is positively associated with its payroll. We investigate this question using a variety of basic statistical techniques to demonstrate how different methods could be employed to study the same issue. We use data from all teams for the 2003 Major League Baseball regular season. To measure "success" we first use the number of wins during the regular season (out of 162 games). An alternate measure of success is whether or not a team qualified for post-season play. Payroll data were obtained from the *USA Today* web site which provides the payroll (in millions of dollars) for each team's entire roster on Opening Day. The data are provided in Table 1 and may be downloaded from [www.amstat.org/publications/stats/data.html](http://www.amstat.org/publications/stats/data.html).

## Analysis #1 – Anecdotes based on the Extremes

The New York Yankees had by far the highest payroll (\$152.75 million) in 2003 and also had the most wins

*Robin Lock is the Jack and Sylvia Burry Professor of Statistics in the Department of Mathematics, Computer Science and Statistics, and the hapless goalie for the faculty/staff ice hockey team at St. Lawrence University.*



Robin Lock

(101) during the regular season. The Detroit Tigers had one of the worst seasons in recent years (only 43 wins) and were among the lowest payrolls (just \$49.17 million). Both of these extreme cases would support an argument for a positive association between success and payroll. On the other hand, two of the eight playoff teams (Oakland and eventual World Series winner Florida) were ranked 23rd and 25th (out of 30 teams) in payroll, while the second-highest payroll (New York Mets) produced a 66-95 record that was the fourth-worst overall. These examples could be used to argue against an association between payroll and success, or even that the association might be negative. The point here is that, even though it's generally a good idea to examine and try to understand the most extreme cases in the data, one should avoid making statistical inferences that are based solely on analyzing the extremes.

## Analysis #2 – Linear Regression and Correlation

Figure 1 shows a scatterplot with a least squares line ( $\hat{w}_{ms} = 66.9 + 0.198 \text{ payroll}$ ) for predicting the number of wins in 2003 from the Opening Day payroll. A *t*-test for whether the slope of this line is positive yields a test statistic of 2.41 and a one-tailed *p*-value of 0.011. Thus, at a 5% significance level, we would find that there is some evidence of a positive linear relationship between payroll and wins. The correlation between wins and payroll is  $r = 0.414$  and would yield the same *p*-value in a *t*-test for positive correlation.

## Analysis #3 – Rank Correlation

One might be concerned about the effects of extreme cases (such as the Yankees' high payroll or the Tigers' low number of wins) on the regression/correlation analysis. To minimize those effects, we compute Spearman's rank correlation using the ranks of both the number of wins and the payrolls. The result is  $r_s = 0.432$  with a *p*-value for a one-tailed test of 0.01 for a sample size of 30. Thus the nonparametric rank procedure agrees well with the more common product-moment correlation analysis.



## Analysis #4 – Difference in Means

An alternate measure of “success” for a team’s season is whether or not they make the playoffs. Just eight MLB teams (six divisional winners and two “wild cards”) out of 30 qualify for post-season play. Do the payroll distributions differ significantly between the teams that make the playoffs and those that do not? The summary data are given in Table 2 and dotplots (with the means) are compared in Figure 2. A one-tailed  $t$ -test to see if the mean payroll for playoff teams is significantly greater than that of non-playoff teams produces a test statistic of 1.37 with a  $p$ -value of 0.10. Thus, at a 5% significance level, we would fail to have sufficient evidence to conclude that the average payroll of playoff teams is significantly greater than the average of teams that missed the playoffs. Although the difference in means, \$18.5 million, is fairly large, the sample sizes, especially of playoff teams, are not very large and the dotplots show a good deal of variability in both groups.

## Analysis #5 – Mann-Whitney Test

Again, if we are concerned about the effects of outliers in the payroll distribution, we could convert to ranks and apply a nonparametric procedure like the Mann-Whitney test to see if the distribution of payrolls for playoff teams is shifted above the distribution of payrolls for non-playoff teams. The result in this case is a  $p$ -value of 0.0986 and a comparable conclusion to the parametric two-sample  $t$ -test.

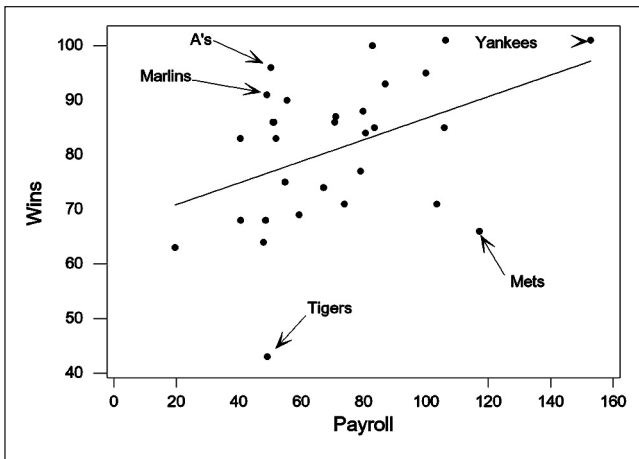


Figure 1: Regression to Predict Wins based on Payroll

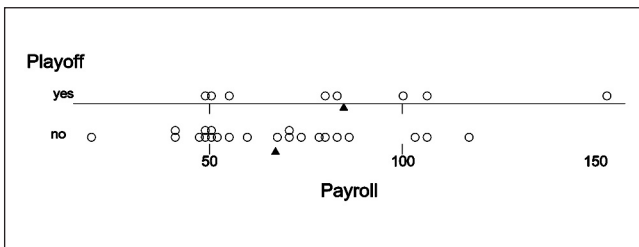


Figure 2: Dotplots of Payrolls for Playoff and Non-Playoff Teams

Teams	Payroll	Wins
New York Yankees*	152.75	101
New York Mets	117.18	66
Atlanta Braves*	106.24	101
Los Angeles Dodgers	105.87	85
Texas Rangers	103.49	71
Boston Red Sox*	99.95	95
Seattle Mariners	86.96	93
St. Louis Cardinals	83.49	85
San Francisco Giants*	82.85	100
Arizona Diamondbacks	80.64	84
Chicago Cubs*	79.87	88
Anaheim Angels	79.03	77
Baltimore Orioles	73.88	71
Houston Astros	71.04	87
Philadelphia Phillies	70.78	86
Colorado Rockies	67.18	74
Cincinnati Reds	59.36	69
Minnesota Twins*	55.51	90
Pittsburgh Pirates	54.81	75
Montreal Expos	51.95	83
Toronto Blue Jays	51.27	86
Chicago White Sox	51.01	86
Oakland Athletics*	50.26	96
Detroit Tigers	49.17	43
Florida Marlins*	49.05	91
Cleveland Indians	48.59	68
San Diego Padres	47.93	64
Milwaukee Brewers	40.63	68
Kansas City Royals	40.52	83
Tampa Bay Devil Rays	19.63	63

\* playoff team

	$n$	Mean	Std. Dev.
Playoff	8	84.6	35.2
Non-Playoff	22	66.1	23.9

## Analysis #6 – Simulating Randomization Tests

We can also assess the significance of the correlation between payroll and wins or the difference in average number of wins between playoff and non-playoff teams by using randomization procedures to build approximate sampling distributions for the statistics in question. Using a software package such as Fathom, we can fix one variable (payroll) and then randomly scramble the values of the other variable (wins or playoff status).

Figure 3 shows histograms of the results of 10,000 randomizations for both the correlation and difference in means. In both cases, we obtain approximate p-values for judging how unusual our original sample is by seeing how often the random scrambles gave a more extreme value for the statistic. Just 93 of the random correlations were bigger or equal to the  $r = 0.414$  that occurred in the actual sample, so we would estimate the p-value for the correlation test to be about 0.0093. For the differences in means, 602 of the randomization differences were \$18.5 million or larger, so the approximate p-value is 0.0602. Each of these p-values is relatively close to the corresponding t-tests discussed earlier, so the conclusions stay the same. There is fairly strong evidence of a positive correlation between number of wins in a season and payroll, but not as convincing evidence that there is a significant difference in average payroll between playoff and non-playoff teams.

## Analysis #7– Logistic Regression

Another way to view the relationship between payroll and success is to build a logistic regression model to predict the probability of making the playoffs based on the Opening Day payroll. We let  $p(x)$  denote the probability that a team makes the playoffs when the payroll is  $x$  and assume a logit model of the form:

$$p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

This is equivalent to assuming that the log-odds of making the playoffs are a linear function of the payroll

$$\ln\left(\frac{p(x)}{1 - p(x)}\right) = \beta_0 + \beta_1 x$$

Using a computer package (Minitab) to obtain maximum likelihood estimates of the parameters  $\beta_0$  and  $\beta_1$  produces the linear fit *log-odds playoffs* = -2.82 + 0.02424*payroll*. Figure 4 shows a plot for the estimated  $p(x)$  as a function of payrolls up to \$200 million. Note that random chance would imply an  $8/30=0.267$  probability for a team to make the playoffs. So the Yankees payroll of \$152.75 million giving an

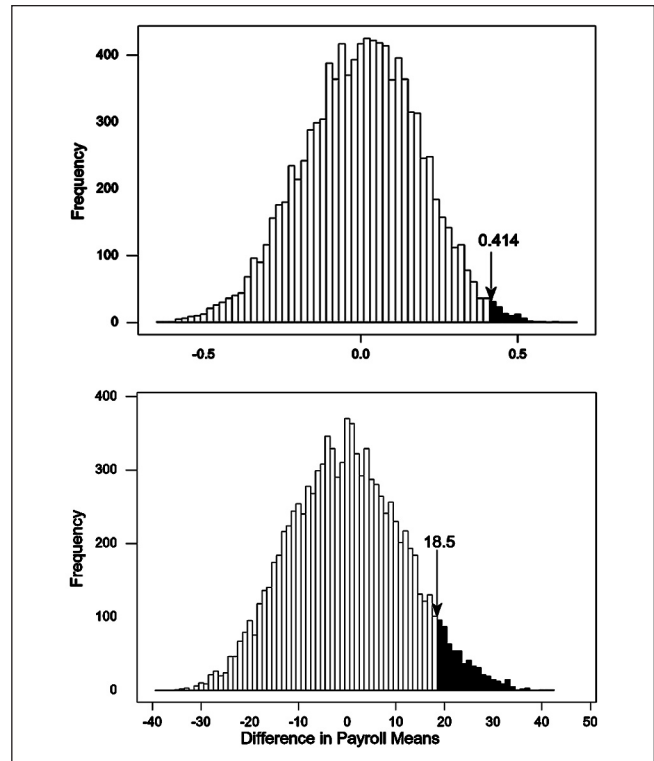


Figure 3: Randomization Results for Correlation and Difference in Means



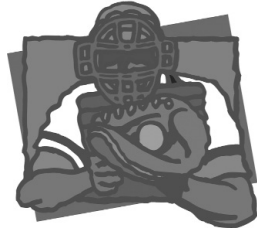
Figure 4: Estimated Probability of Making the Playoffs: Logistic Fit based on Payroll

estimated probability of making the playoffs at 0.73 would seem to be pretty dramatic. By contrast, the Devil Rays \$19.6 million payroll yields just a 0.09 probability of making the playoffs. A test that the slope coefficient  $\beta_1$  is positive yields a p-value of 0.063, so the evidence is a bit stronger than the test for a difference in mean payroll based on playoff participation, but not as strong as the procedures that used wins as the measure of success.

## Suggestions for Additional Investigations

The data presented in this article only used results from the 2003 season. Is this really a sample from some

population or are we actually using the entire population, in which case statistical inference would not be required? In the work we presented, we are assuming that the results for 2003 are a convenient sample of wins and payrolls for a "typical"



MLB season—with the advantage of being able to compare the salaries in constant dollars. One might expand the data to include other seasons, provided the payrolls from different years could be adjusted for baseball-specific inflation. For example, it would be reasonable to use the payroll ranks from different years or convert the payrolls within a year to z-scores using the mean and standard deviation for payroll values within that year.

One might question the use of Opening Day payroll as the indicator of a team's spending habits. Highly paid stars may be injured for large parts of a season or a struggling team might reduce its payroll by trading a player with a big contract to a pennant contender. Perhaps the total salaries of players on the rosters at some other point of the season, such as the August 31 deadline for acquiring playoff eligible players, would be a better payroll measure.

Of course, it would also be interesting to use these techniques to explore similar trends in other sports such as such as professional hockey, football, or basketball.

## Conclusion

The data show some support for the notion that more successful teams tend to have higher payrolls, although the evidence is less convincing when "success" is measured by a gross categorical variable (playoff participation), than when using a more refined quantitative variable (number of regular season wins). However, current (and future) team owners and general managers should note that this is an observational study from a single season. As any good statistics student knows, we should be reluctant to infer a cause/effect relationship in such a situation. Spending more money on player salaries alone will never guarantee success, whether measured by number of wins or chances of making the playoffs.

## References

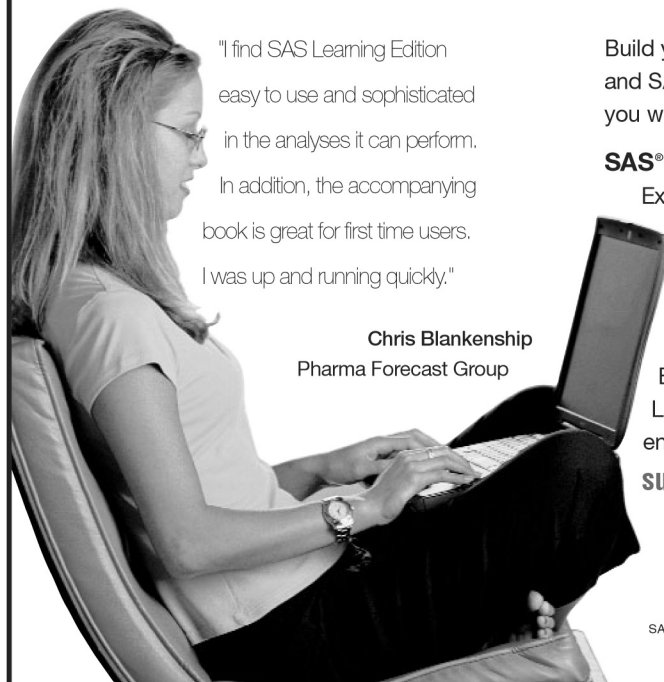
Lewis, M. (2003), *Moneyball: The Art of Winning an Unfair Game*. New York: W.W. Norton & Company.

## Web Resources

*USA Today's Baseball Salary Database*, salary information for teams and individual players since 1988, <http://asp.usatoday.com/sports/baseball/salaries/default.aspx>. ■

## Professional development leads to professional advancement...

### Invest in yourself!



"I find SAS Learning Edition easy to use and sophisticated in the analyses it can perform. In addition, the accompanying book is great for first time users. I was up and running quickly."

Chris Blankenship  
Pharma Forecast Group

Build your SAS skills and achieve greater success with SAS Learning Edition and SAS Self-Paced e-Learning. Work at your own pace whenever and wherever you want.

#### SAS® Learning Edition

Explore the powerful functionality of SAS with instruction paced for the beginner. SAS Learning Edition provides a low-cost, PC-based entry point to SAS' new generation of business intelligence software—all on a single CD-ROM. [support.sas.com/le](http://support.sas.com/le)

#### SAS® Self-Paced e-Learning

Experience the convenience of desktop training designed by SAS experts. License and gain immediate access to Web-based lessons, courses, or entire libraries to meet your individual needs—all at an affordable price.

[support.sas.com/selfpaced](http://support.sas.com/selfpaced)

*The Power to Know.*



SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © Indicates USA registration. © 2003 SAS Institute Inc. All rights reserved. 260214US.1203

# A Simulation Alternative to the Two Sample $t$ -test



**Josh Tabor**

In studying “The Antihypertensive Effects of Fish Oil,” researchers randomly divided 14 male volunteers with high blood pressure to one of two treatments. The first treatment was a four-week diet that included fish oil and the second was a four-week diet that included regular oil. The response variable was the reduction in diastolic blood pressure (mm of mercury). The results of this study are shown below (Knapp and FitzGerald (1989), cited in Ramsey and Schafer (2002), page 23).

Fish Oil:	8	12	10	14	2	0	0
Regular Oil:	-6	0	1	2	3	-4	2

The null hypothesis for this study is that the population mean reduction in diastolic blood pressure is the same for both treatments ( $H_0: \mu_{fish\ oil} - \mu_{regular\ oil} = 0$ ) and the alternative hypothesis is that the population mean reduction in diastolic blood pressure with the fish oil diet is greater than the mean reduction in diastolic blood pressure with the regular oil diet ( $H_a: \mu_{fish\ oil} - \mu_{regular\ oil} > 0$ ).

In the experiment, the sample mean reduction for the fish oil group was  $\bar{x}_{fish\ oil} = 6.57$  and the sample mean reduction for the regular oil group was  $\bar{x}_{regular\ oil} = -1.14$ . Thus, since the difference in the treatment means was  $= 7.71$ , it *appears* that the fish oil treatment did help reduce diastolic blood pressure for the men in the study.

Of course, there are two possible explanations for the observed difference of 7.71. The first is that the null hypothesis is wrong and the fish oil actually does a better job in reducing blood pressure than the regular oil. The second explanation is that the null hypothesis is true and that the observed difference was due to the randomization of subjects to treatments. Since the blood pressure for each subject has natural variability, different randomizations will produce different results, even when all experimental conditions are the same. Thus, the difference of 7.71 could have been due to the particular men that ended up in each group and not due to any difference in treatments.

How can we decide which explanation is correct? The traditional method is to compute a  $p$ -value using a two sample  $t$ -test. However, an alternative approach is to estimate the  $p$ -value with a simulation. In both procedures, we begin by assuming the null hypothesis

is true. That is, the reduction in mean diastolic blood pressure is the same for men with the fish oil diet and men with the regular oil diet. For example, if the null hypothesis is true, then the first subject in the table above would have had a reduction of approximately eight no matter which group he was assigned to.

Thus, to simulate the results of this experiment under the hypothesis of no treatment effects, we can write each of the 14 data values on a separate slip of paper and mix them up in a hat. Then, we could randomly select seven reduction values to go into the fish oil group and put the remaining seven reductions into the regular oil group.

One possible randomization could be:

Fish Oil:	-1	10	-3	0	14	2	2
Regular Oil:	-6	-4	0	12	8	0	2

In this trial,  $\bar{x}_{fish\ oil} - \bar{x}_{regular\ oil} = 3.71 - 1.71 = 2.00$ .

Another possible randomization could be:

Fish Oil:	12	-6	0	2	0	2	0
Regular Oil:	1	-3	8	2	14	-4	10

In this trial,  $\bar{x}_{fish\ oil} - \bar{x}_{regular\ oil} = 1.43 - 4 = -2.57$ .

Other trials would give different allocations of subjects and differences in sample means. Ideally, we would look at every possible randomization to calculate the exact probability of getting a result of 7.71 or above. However, it would be cumbersome to analyze all 3,432 different randomizations and calculate all the possible differences in sample means (making sure not to repeat any randomizations). Instead, we will use technology to do 1,000 trials without worrying about repeats. The distribution of the differences in sample means from this simulation is shown in Figure 1.

As we can see, it is very unlikely to get a difference of 7.71 or above by random chance (the approximate  $p$ -value from this simulation is .007). Thus, it appears that fish oil is more effective than standard oil in reducing diastolic blood pressure.

Using the traditional  $t$ -test approach, we get a  $p$ -value of .0065, which is very similar to the value in our simulation. Since the  $t$ -test requires only a few keystrokes or mouse clicks and gives similar results,

*Josh Tabor is an AP Statistics teacher at Wilson High School in Hacienda Heights, California.*

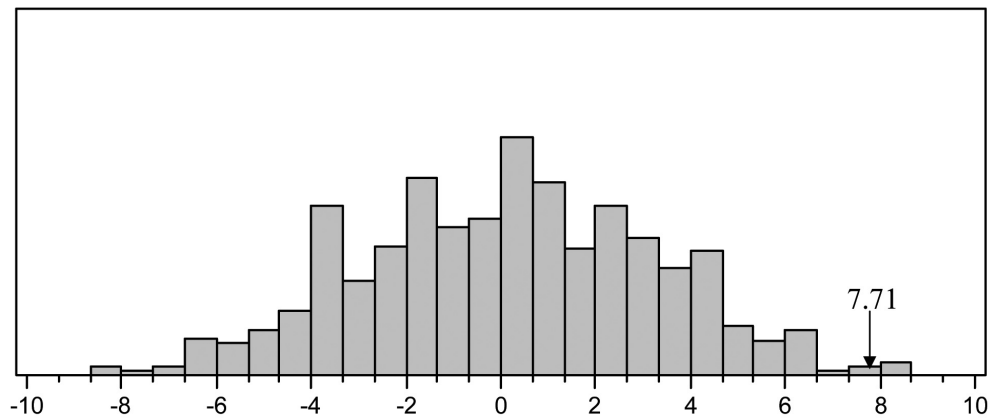


Figure 1: Distribution of differences in sample means

why should we consider using the simulation approach? First, a  $t$ -test requires either that the observations come from a normal population or that the sample sizes be large. The simulation method is not restricted by either condition. Second, the simulation approach generates a concrete sampling distribution which makes it easy to see where the observed value falls in the distribution without relying on any formulas. Third, the simulation method makes it clear that the random assignment of subjects to treatment groups is what drives the inference procedure.

Although this article presented a simulation approach to a two sample  $t$ -test, this type of randomization procedure can be used to estimate the  $p$ -value for any common significance test—and the

best part is that you don't need to remember any formulas!

I would like to thank Dick Schaeffer and Kim Robinson for their help in developing this article.

## References

Knapp, H.R., and FitzGerald, G.A. (1989), "The Antihypertensive Effects of Fish Oil. A Controlled Study of Polyunsaturated Fatty Acid Supplements in Essential Hypertension," *New England Journal of Medicine*, 320: 1037-1043.

Ramsey, F. and Schafer, D. (2002), *The Statistical Sleuth (2nd ed.)*, Belmont, CA: Duxbury. ■



# Are You a Student Majoring in Statistics?

You should become a student member of the American Statistical Association! For only \$10 a year, you can join the premier statistical organization in the United States. With your membership, you will receive member discounts on all meetings and publications, access to job listings and career advice, as well as networking opportunities to increase your knowledge and start planning for your future in statistics.

## Join Now!

To request a membership guide and an application, call 1-888-231-3473 or join online now at [www.amstat.org/join.html](http://www.amstat.org/join.html).



# Review of *A Life of Sir Francis Galton: From African Exploration to the Birth of Eugenics*

Gillham, Nicholas Wright. Oxford University Press. 2001. ISBN 0-19-514365-5.



Chris Olsen

In the interest of fair disclosure, I need to say right up front that I am not a particular fan of biographies. Part of this has to do with my rebellious nature. My English teachers long ago thought it was terribly important to read Boswell's *Life of Johnson* (or Johnson's *Life of Boswell*, or maybe they coauthored a biography of someone else?). Neither of them had, so far as I could tell, proved even one theorem or worked out the sampling distribution of a single statistic. Ever the optimist, I kept reading whoever's biography that was until the very last page before discovering this lack of mathematics, and my disappointment was rampant. Ever since, biographies have been on my short list of genres to avoid.

The second reason I am not a particular fan of biographies is what might be referred to as the "Lost-in-Leviticus" phenomenon. As may be recalled, somewhere in the first five books of the Bible everybody starts to beget everyone else, resulting in a long line of characters that seem irrelevant to the story, except that many of them were lawyers and they passed a lot of laws.

Be that as it may, I would like to recommend a biography that you can get in plenty of time for reading over spring break: *A Life of Sir Francis Galton: From African Exploration to the Birth of Eugenics*, by Nicholas Gillham (SFG). As all students of statistics and especially readers of Stephen Stigler's work know, Galton was present at the creation of statistics, virtually single-handedly inventing correlation and regression. Thus, this book actually has topics of interest in it—mathematics and statistics. Secondly, while there is a certain amount of ancestral review, the begetting is kept to a minimum. In Galton's case, at least the ancestors are interesting! We learn, for example, that Galton was the grandson and cousin of Darwins, the most famous being his cousin, Charles, of *Origin of Species* fame. Grandfather Darwin, Erasmus by name,

was a physician, inventor, and poet—just the sort of individual my English teachers would have liked. Not only is the begetting kept to a minimum, but there is only one potential begetting in the whole book and it is only spoken of in the most delicately Victorian tones. Modern literature is apparently required to delve into intimate details about such things, but in SFG we find only that Galton may have dallied with an individual referred to as "a woman of easy virtue."

The most refreshing aspect of SFG is that it is written by a nonstatistician, mostly for nonstatisticians. Now, there is nothing wrong with statisticians writing for statisticians, but statisticians tend to focus in some detail on all those statistical things, at the expense of, well, all those other things. And in SFG, there are a lot of other things! Sir Francis was a very busy fellow beyond "merely" inventing correlation and regression.

Statistical types are probably aware, with some chagrin, that Galton was very involved in eugenics; in fact, Galton actually coined the word, deriving it from Greek parentage, meaning "good in birth." The sorry history of the eugenics movement in the 20th century casts a less than favorable light on this aspect of Galton's interests, but I felt the treatment of this topic provided a proper, that is, 19th century, perspective. Upon reading Darwin, Galton concluded that it might be possible to improve mankind by selective breeding, an idea that seemed more plausible in a time of Victorian science and British class society than it does today after the horrible excesses of Nazi Germany. Gillham writes that Galton meant well in his efforts to improve mankind, but "he viewed the world through the lens of class, privilege, and the predominant role played by men in virtually all affairs in Victorian England." While not mincing words about eugenics, Gillham provides a social and scientific historical context to Galton's thought that will broaden one's perspective about this aspect of Galton's work. Gillham is a geneticist and professor of biology, so his discussion of these issues and of Galton's understanding of the genetic science is wonderfully authoritative.

---

*Chris Olsen teaches mathematics and statistics at George Washington High School in Cedar Rapids, Iowa. He has been teaching statistics in high school for 25 years, and AP Statistics since its inception.*

Gillham's writing brings to this work something very rare for historical biographies: relevance for modern times. We live in an exciting time, in which the human genome has been mapped, and ethicists have begun to worry anew about manipulating mankind's hereditary destiny. In the third section of SFG, "The Triumph of the Pedigree," Gillham discusses the development of Galton's statistical methods in the light of his investigations of hereditary mechanisms, psychology, fingerprinting (!), and biometry in general. If you are interested in science, statistics, or the history of scientific method and thought, this section of the book is very illuminating.

What is also illuminating is what we might call the First Half of Galton's career—Galton was an accomplished explorer and geographer early on, mounting and financing his own expedition into Africa, and was elected to the Council of the Royal Geographic Society. He was an accomplished writer about travel, as well as a meteorologist of some

renown. In this explorer/writer/geographer role he even carried on a feud reminiscent of the Pearson-Fisher relationship. It seems that in his travel writings, Galton ran afoul of Henry Stanley, of "Dr. Livingstone, I presume" fame. Galton suggested that perhaps Stanley was engaging in sensationalism rather than searching for the good Dr. L. The story of Galton's expedition is recounted, and one does not notice the absence of equations and diagrams.

Finally, I'm afraid I'm going to have to eat some crow here. Gillham's biography of Galton is a page-turner and a delight to read. Gillham weaves statistics, science, and society into a coherent story; he is detailed without being mind-numbing, and always has an eye on the "big picture." His writing flows extremely well, and his narration always keeps the attention of the reader.

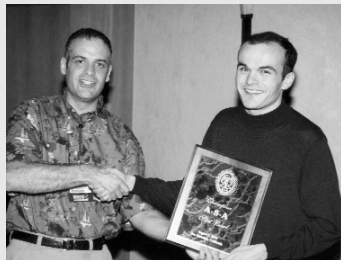
In short, this is a biography even an English teacher would love. ■

## Advice from the 2003 Stat Bowl Champion

By Wesley Eddings, Kenyon College

**D**o you watch television game shows religiously? Does the new edition of Trivial Pursuit appear on your holiday gift wish list every year? Did you never outgrow your childish impulse to show off in front of a roomful of strangers? Or would you just like \$500 to attend this year's Joint Statistical Meetings in Toronto? If you can answer any of these questions in the affirmative, you're a perfect candidate for this year's ASA Stat Bowl!

I volunteered for the 2003 Stat Bowl in San Francisco after I saw an announcement in *Amstat News*. On the



Mark Payton and 2003 Stat Bowl champion Wes Eddings.

Tuesday of the meeting I survived three rounds of grueling competition to emerge as champion, with the highlight a close final victory over David Hitchcock of the University of Florida (and author of a published paper on the history of the Metropolis-Hastings

algorithm). I received a nice plaque (now a conversation piece in my office at Kenyon College) and had a great time in once again indulging my fondness for trivia. I even won a better prize—\$500 in travel reimbursements—than I won when I appeared on *Jeopardy!* three years ago!

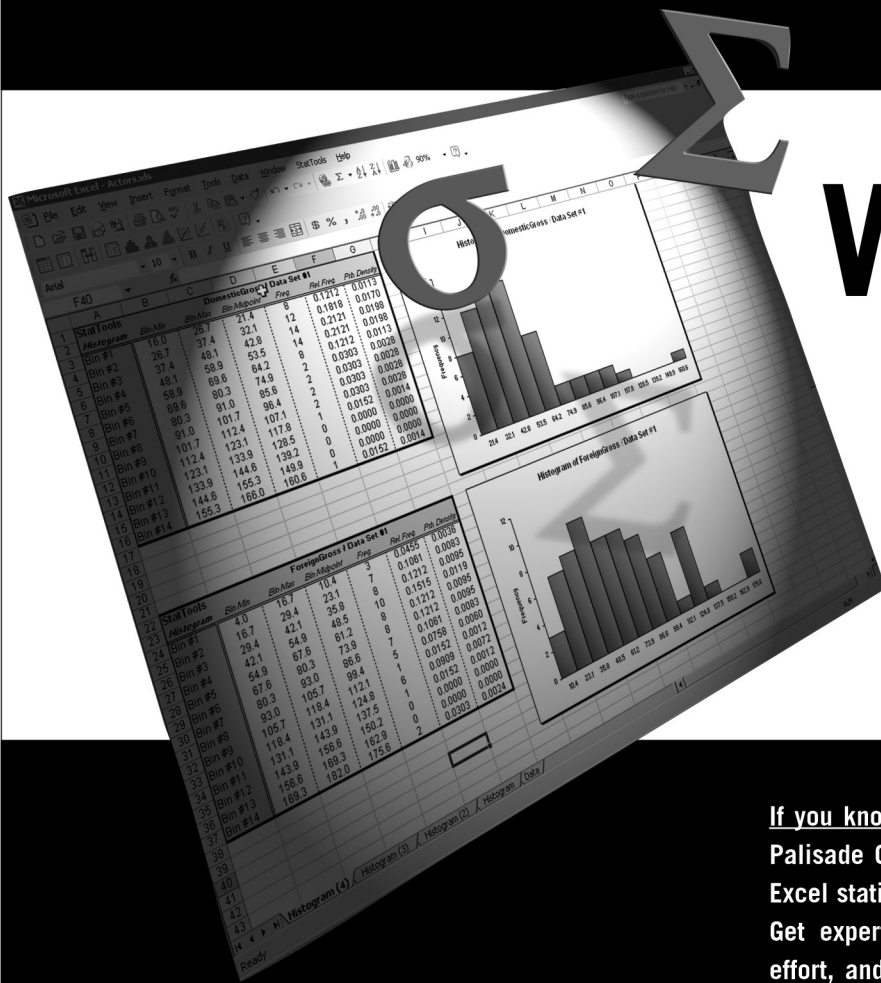
I didn't really study for the competition as such; my coursework and Ph.D. dissertation in biostatistics at Johns Hopkins served as an indirect form of study. My research area—foundational and philosophical topics in statistics—may have been an advantage, as I read a wide variety of older statistical papers and books that play a

lesser role in most modern research. I did relatively well on questions about the statistical literature and famous statisticians and relatively poorly on facts and problems from the mathematical theory of statistics. I was fortunate that none of my course instructors were in the audience to witness the deterioration of my knowledge of the gamma distribution, moments, and order statistics! Questions about classical mathematical statistics were not uncommon, but general questions about statistical ideas, people, and history made up most of the competition. My notes show questions on naming authors of papers and books from their titles, naming universities at which famous statisticians concluded their careers, and (the horror!) answering with statisticians' middle names when prompted with their first and last names and middle initials! (I don't regret being in the audience for that one.) Connoisseurs of statistical puns shouldn't miss this year's competition.

If you're fortunate enough to participate in the 2004 Stat Bowl, don't spend much time studying "factoids" specifically for the competition. Facts crammed at the last minute don't sink sufficiently deeply into your memory, and you probably won't be able to recall them within the time limits of the game. Reading a good book about our subject is a better way to prepare. Be sure to look over the rules in advance of the competition, and pay attention in the early rounds: questions in later rounds in 2003 were often quite similar to those in earlier rounds. And take your time in any questions involving calculation, so that you don't repeat my mistake of summing 1, 4, 9, and 16 and arriving at 25 as the answer.

Have a great time and bring a cheering section of your classmates, and maybe I'll see you there! ■

# StatTools Excels at Statistics.



## With No Learning Curve.

If you know Excel, you can use StatTools. StatTools from Palisade Corporation is a new Excel add-in that replaces Excel statistics functions with an accurate, proven toolset. Get expert statistics where you need them. Save time, effort, and money. And you can easily create customized statistical procedures using Excel's built-in VBA language.

Download a free trial version at [www.palisade.com](http://www.palisade.com) today.

Order StatTools now and save \$250!

Use code ST0104 at [www.palisade.com/stattools/](http://www.palisade.com/stattools/) or when calling 800-432-RISK. Be sure to ask about special academic pricing.

 **PALISADE**  
DECISIONS WITH VISION