STATS

THE MAGAZINE FOR STUDENTS OF STATISTICS

FALL 2000 • ISSUE # 29

INSIDE:

We've Learned Inference About the Mean; What Should We Learn Next?

Start an IV – STAT!

A Day in the Life of a Statistical Manager at an Aerospace Company

Major League Soccer: Predicting Attendance for Future Expansion Teams

Statistical Computing Section
American Statistical Association
Student Paper Competition 2001

The Statistical Computing Section of the ASA is sponsoring a student paper competition on the topic of Statistical Computing. Students are encouraged to submit a paper in this area, which might be original methodological research in statistical computing, some novel computing application in statistics, or any other suitable contribution (for example, a software-related project). The selected winners will present their papers in a topic contributed session at the 2001 Joint Statistical Meetings in Atlanta. The Section will pay registration fees for the winners as well as a substantial allowance for transportation to the meetings and lodging (which in most cases covers these expenses completely).

Anyone who is a student in the fall of 2000 (undergraduate, Masters, or Ph.D.) is eligible to participate. An entry must include an abstract, a six page manuscript, a resume, and a letter from a faculty member familiar with the student's work. The manuscript should be single-spaced in a 10 point font with one inch margins (this is consistent with ASA Proceedings guidelines). All figures, tables and references must be included in the six-page limit. The applicant must be the first author of the paper. The faculty letter must include a verification of the applicant's student status and, in the case of joint authorship, should indicate what fraction of the contribution is attributable to the applicant. We prefer that electronic submissions of papers be in Postscript or PDF. All materials must be in English.

All application materials MUST BE RECEIVED by 5:00 PM PST, Wednesday January 3, 2001 at the address below. They will be reviewed by the Student Paper Competition Award committee of the Statistical Computing Section. Selection will be based on a variety of criteria at the discretion of the selection committee, and will include innovation and significance of contribution, amongst others. Award announcements will be made in late January, 2001. The decision of the selection committee will be final.

Information on the competition can also be accessed on the website of the Statistical Computing Section. A current pointer to the website is available from the ASA website at www.amstat.org. Inquiries and application materials should be emailed or mailed to:

Student Paper Competition
c/o Lionel Galway
The RAND Corporation
1700 Main St.
Santa Monica, CA 90407
Lionel_Galway@rand.org

STATS

STATS

STATS

The Magazine For Students Of Statistics

Fall 2000 • Number 29

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, Virginia 22314-3415 USA; (703) 684-1221; fax: (703) 684-2036; Web site: *www.amstat.org*

STATS is published for beginning statisticians, including undergraduate and graduate students who have a special interest in statistics, and is distributed to student members of 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 Editor, Jerome P. Keating. Addresses of the Editor and Editorial Board are listed above.

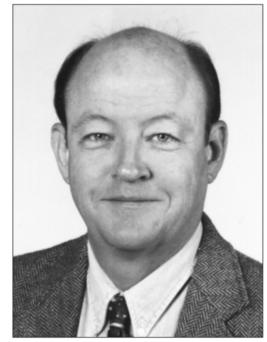
Requests for membership information, advertising rates and deadlines, subscriptions, and general correspondence should be addressed to *STATS* at the ASA office.

Produced by Oakland Street Publishing • Printed in U.S.A.

Features

Departments

Editor's Column



Jerome P. Keating

I want to welcome students and faculty back to a new academic year and all readers to a new year of *STATS*. This issue engages some new themes and continues to embrace many of the time-treasured ones of our profession.

■ Feature Article

We are privileged to have a thought-provoking article on statistical inference from one of the doyens of our discipline, Albert Madansky, Professor Emeritus at the University of Chicago. He discusses the issue of parallel concepts to the Central Limit Theorem but for statistics other than the sample mean. By way of example, he reminds students of the folly of looking at density estimates alone to check for normality.

I strongly recommend his monograph, *Prescriptions for Working Statisticians*, for everyone who has worried about checking assumptions in ANOVA's or regression.

■ Column Articles

In this issue, we present a student project by Diego De Rose and Rafael Galarza on "Estimating Soccer Attendance." They present a regression analysis to estimate soccer attendance for future expansion teams of the MLS. These Hispanic students are evidence that the emerging trend of assimilation and diversity is well underway in American universities.

In the column "Student Voices," Amanda Elizabeth Brooks, a freshman at the University of North Carolina at Chapel Hill, gives us a glimpse at the statistics associated with emergency medical care. She describes how she merged her volunteer work as an Emergency Medical Technician with her high school AP Statistics class. In this article she analyzes the success rate of intravenous cannulation in emergency situations.

Eric P. Fox of Pratt & Whitney discusses the frantic life of a statistician in the aerospace industry. He uses a wide spectrum of methods, including Design of Experiments (DOX), logistic regression, product reliability, manufacturing quality control, and more. Eric overlays the technical nature of his job with a time schedule that demands a considerable amount of travel. I am sure that you will find this to be a very stimulating article.

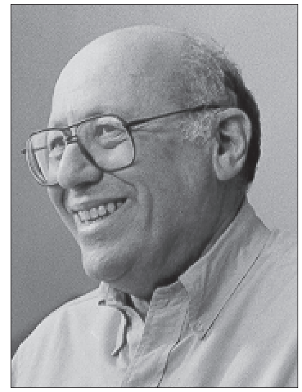
David Drain, Intel Corporation, provides us with insight into the use of statistical methods in the manufacture of computer chips and boards. He focuses on the use of Statistical Process Control (SPC) and Design of Experiments. He provides the reader with a sketch of his progression up the company ladder over the last 13 years and he recommends several useful books for those interested in pursuing a career in the industry.

Terry Allen from the State of Utah discusses the interesting job of being a medicaid fraud statistician. He describes how statisticians can uncover fraudulent billing practices by physicians. These statistical methods are used by governmental agencies to insure compliance with state and federal regulations.

Stephen Gulyas of the Pfizer Research Center provides you with a day in the life of a pharmaceutical industry statistician. Follow Steve's typical day at Pfizer from 8:07 AM until 5:47 PM. He provides you with some very important history into the origins of statistics in the pharmaceutical industry through the Kefauver-Harris amendments passed by Congress in 1962. If you attended the JSM in Indianapolis, you were impressed with the number of pharmaceutical companies that were hiring statistics graduates.

The "Outlier…s" column is back with several new and challenging exercises. Check Allan J. Rossman's sources on ranking colleges, places to live, books, presidents, and vacation spots. Can you replicate the rankings for places to live?

We've Learned Inference About the Mean; What Should We Learn Next?



Albert Madansky

In the very first course in Statistics, the standard situation presented is one where the data are a set of n independent observations from some (usually unspecified) distribution. The student is then taught to make inferences about the distribution's mean μ based on the sample mean and either the population standard deviation σ or the sample standard deviation s. The student learns that the inferential procedure is based on a broadly-applicable theorem, the Central Limit Theorem, which says that the sampling distribution of the sample mean is approximately normal, with mean μ and standard deviation, $\sigma/\sqrt{n}$, almost regardless of the nature and shape of the distribution from which the data were drawn.

It is rare, though, that the student is taught how to verify the independence assumption. Usually this matter is deferred until the student has taken a course in Regression Analysis and been exposed to the concept of autocorrelation. (To be fair, some first courses in Statistics do introduce so-called nonparametric or distribution-free tests for independence, such as variants of the runs test. These tools have the virtue of being easy to learn and implement and, by their nature, applicable to data drawn from almost any distribution. But by their nature they have very little power.) Unfortunately, the validity of the standard procedures for verifying independence based on autocorrelations requires that the data being autocorrelated are drawn from a normal distribution—a requirement that is surely not met by every data set. After this long wait the student finds that the procedure he has been taught for verifying independence is of only limited value.

If the introductory course gets beyond inferences about μ, it usually focuses on inferences about the variance σ^2. Unfortunately, there is no equivalent to the Central Limit Theorem at work here so that our inferences can be made for a variety of distributions from which data are drawn. The student is taught that, if the data are independent and drawn from a normal distribution, then the sampling distribution of $(n-1)s^2/\sigma^2$ is chi-squared (χ^2) with $n-1$ degrees of freedom, i.e.,

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1).$$

Again this is a requirement not met by every data set.

Some introductory courses teach the student how to make inferences about some other parameter, such as the population median ν or the population standard deviation σ, from its sample counterpart. The student is told that the *"large-sample"* sampling distributions of these statistics are normal, with mean equal to the corresponding population parameter, almost regardless of the distribution of the data from which the samples were drawn. But the student is also told that the variances of these normal sampling distributions are somewhat more complex than the easy-to-remember σ^2/n.

In the case of the sample median, the variance of its sampling distribution is

$$\frac{1}{4nf_x^2(\nu)}$$

and depends on knowledge of the density function evaluated at the population median, ν.

In the case of the sample standard deviation,

Albert Madansky is the H.G.B. Alexander Professor Emeritus of Business Administration in the University of Chicago's Graduate School of Business, and is a Fellow of the ASA, IMS, and Econometric Society. For the past 45 years he has developed statistical and econometric methods and applied them in marketing, finance, and forensics. This paper is an adaptation of Chapter 0 of his book, Prescriptions for Working Statisticians.

the variance of its sampling distribution is

$$\frac{\mu_4 - \sigma^4}{4\sigma^2(n-1)}$$

where μ_4 is the population fourth central moment.

For pedagogical ease in teaching inference about these parameters, the course restricts itself to making a particular assumption about the distribution from which the data are drawn, sometimes the uniform but usually the normal distribution.

Uniform Distribution: The uniform distribution over the interval $[\alpha, \beta]$ is chosen because $f_x(\nu)$ is of the simple form $f_x(\nu) = 1/(\beta - \alpha)$, so that

$$\frac{1}{4nf_x^2(\nu)} = \frac{(\beta-\alpha)^2}{4n}.$$

Since $\sigma^2 = (\beta - \alpha)^2 / 12$ and $\mu_4 = (\beta - \alpha)^4 / 80$, the quantity

$$\frac{\mu_4 - \sigma^4}{4\sigma^2(n-1)} = \frac{(\beta-\alpha)^2}{60(n-1)}.$$

Normal Distribution: If the normal distribution is chosen then $f_x(\nu) = 1 / (\sqrt{2\pi\sigma})$, so

$$\frac{1}{4nf_x^2(\nu)} = \frac{\pi\sigma^2}{2n}$$

and, since $\mu_4 = 3\sigma^4$, the variance takes on the simple form:

$$\frac{\mu_4 - \sigma^4}{4\sigma^2(n-1)} = \frac{\sigma^2}{2(n-1)}.$$

In the latter case we must check whether or not the data came from a normal distribution.

With respect to checking the assumption that the data came from a normal distribution, the student is told a few salient facts about the percent of observations encompassed within intervals of the form $\mu \pm k\sigma$, where k = 1, 2, and 3, and asked to look at a histogram of the sample data and compare it visually with the familiar bell-shaped curve. If the histogram looks bell-shaped and if the fractions of the data falling within 1, 2, and 3 sample standard deviation from the sample mean are approximately 67%, 95%, and 99%, then they should feel comfortable with behaving as if the data came from a normal distribution.

Let's look at some data to see whether this prescription helps. Table 1 gives the sixth grade verbal achievement scores for 20 randomly chosen schools (see Table 14-1 of Mosteller, Fienberg, and Rourke [1983]).

We find that the sample mean is 35.0825, the sample standard deviation is 5.81705, and following is a histogram of the 20 scores (Fig. 1):

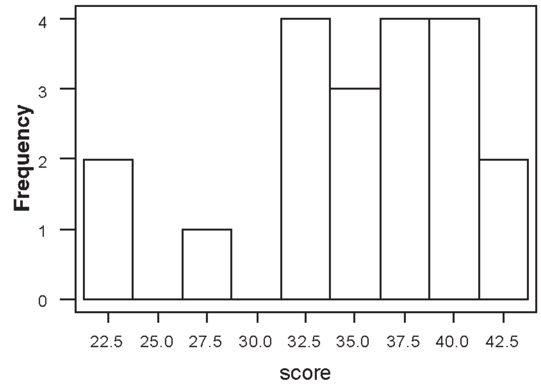


Figure 1

Score Midpoint	Count	Percentage	Cumulative percentage
23.00	2	10.0	10.0
26.00	1	5.0	15.0
29.00	0	0.0	15.0
32.00	4	20.0	35.0
35.00	3	15.0	50.0
38.00	4	20.0	70.0
41.00	5	25.0	95.0
44.00	1	5.0	100.0
Total	20	100.0	

The histogram does not appear very bell-shaped. Should we therefore assume nonnormality of the scores? Let's look at a frequency distribution of the data:

We note that 85% of the sample, i.e., 17 observations, lie in the interval 35.0525 ± 5.81705, that is, within one (sample) standard deviation of the sample mean, in contrast to the 67% expected. There are 19 observations, or 95% of the sample, within two standard deviations, and all 20 observations are within three standard deviations of the sample mean. We are at a loss as to what to make of this possibly conflicting visual evidence about the normality of the distribution of these scores.

Exploratory Data Analysis (EDA) suggests that we look at a plot of the data relative to its median,

Table 1

37.01	26.51	36.51	40.70	37.10	33.90	41.80	33.40	41.01	37.20
23.30	35.20	34.90	33.10	22.70	39.70	31.80	31.70	43.10	41.01

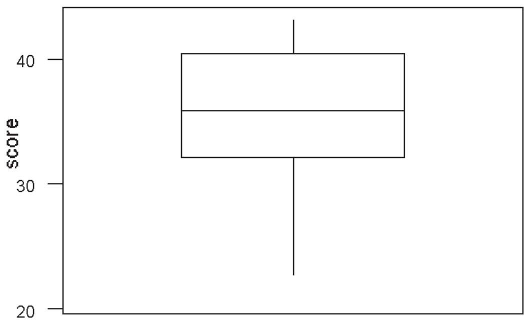


Figure 2

35.86, and its quartiles, 32.13 and 40.45, via a boxplot (Fig. 2). The boxplot highlights any asymmetry in the data, as well as the "fatness" of the tails of the distribution.

Asymmetry does not appear to be a problem, but the left tail looks a lot longer than the right tail (distance from min to first quartile is

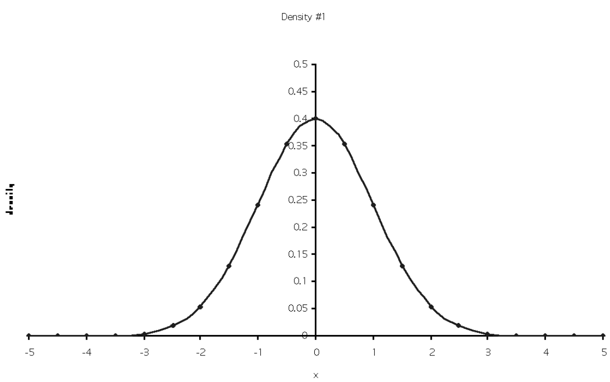


Figure 3

32.13 – 22.70 = 9.43; distance from third quartile to max is 43.10 –40.45 = 2.65). In short, we are left in a quandary about the validity of assuming

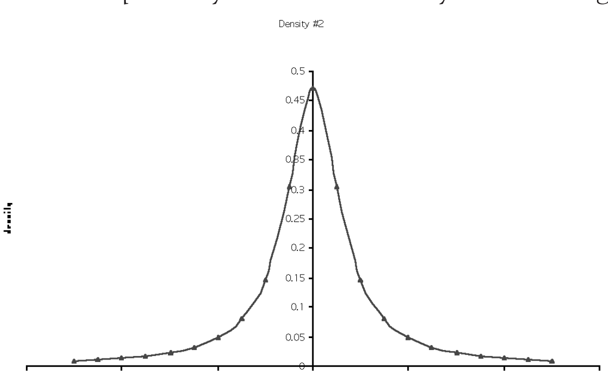


Figure 4

normality in analyzing these data.

I always show my introductory Statistics

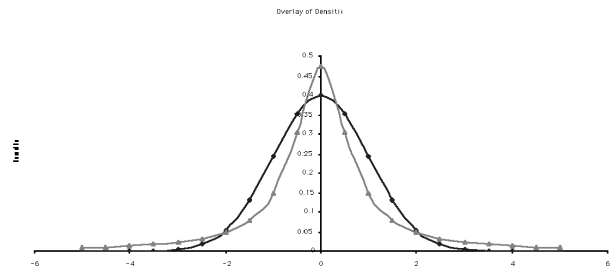


Figure 5

students the two density functions (Figs. 3 and 4) and ask them which portrays a normal distribution. (They are allowed to answer "Both.")

Sometimes I show them the two densities superimposed on each other (Fig. 5) and ask the same question.

The most common answer is "Both." When I then tell them that one is and one isn't a normal distribution and ask them which is the normal distribution, about half the class chooses Fig. 4. I then tell them that Fig. 3 is a normal distribution with $\mu = 0$ and $\sigma = 1$ and Fig. 4 is a Cauchy distribution with 0 as its median and .6745 as its semi-interquartile range (the same as that of a standard normal distribution).

I sometimes make life even harder, by showing them the superimposed densities (Fig. 6), one

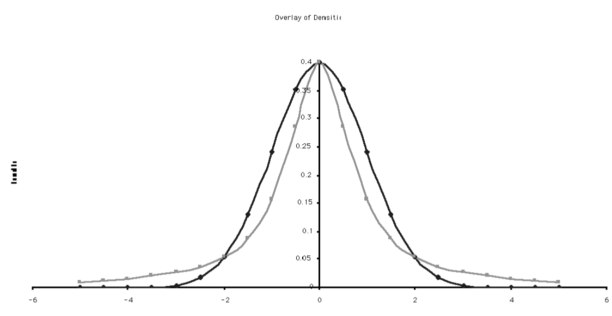


Figure 6

being the normal distribution with $\mu = 0$ and $\sigma = 1$ and the other being a Cauchy distribution with 0 as its median and $\sqrt{2/\pi}$ = .79788 as its semi-interquartile range (so that the two densities' functions have the same value at x = 0).

The point to this exercise is to impress upon the students that their visual intuition about the shape of normal density functions is untrained, and so relying on visual inspection of histograms to determine whether the underlying data were from a normal distribution is unreliable.

With this, the notion of the normal probability plot is introduced. The motivation is that one needs to compare the data at hand with a

Table 2

0.18593	-1.12690	0.06165	0.74198	0.31325	0.31325	1.40377
-0.44602	1.01639	0.44602	-1.40377	-0.06165	-0.18593	-0.58740
-1.87129	0.58740	-0.74198	-0.91718	1.87129	1.01639	

"reference set," i.e., what a "perfect" set of n observations from a normal distribution would look like. I begin with n=1, ask what one would "expect" that one observation to be, and invariably get the answer μ, as the student recognizes that I am asking for E[X]. I next ask: What one would "expect" if n were 2? It takes a bit of discussion to convince the student that we would expect the two observations to be the 1/3 point and 2/3 point of the cumulative normal distribution (i.e., $\mu - .4307\sigma$ and $\mu + .4307\sigma$), and that when n = 1 we would expect the ½ point of the cumulative normal distribution, i.e., the median μ. Ultimately the student recognizes that in general the n "expected" observations are those corresponding to

$$\mu + F^{-1}(i / (n + 1))\sigma \text{ , for } i = 1,\ldots,n,$$

where F is the cumulative normal distribution function. The "expected" observations from a standard normal distribution are sometimes called "normal scores," and computer packages such as MINITAB routinely generate them as reference sets to use in a comparison with the observations at

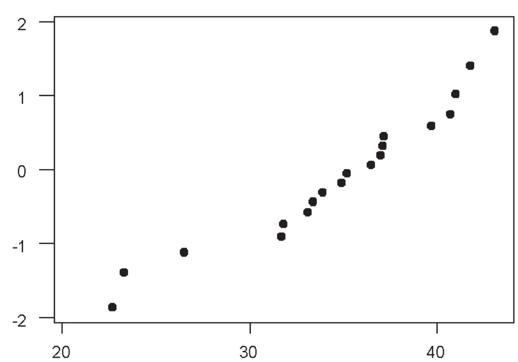


Figure 7

hand.

For example, the normal scores associated with the data of Table 1 are shown above (Table 2):

One could plot the data versus these normal scores to produce Figure 7.

One can also estimate the data's reference set

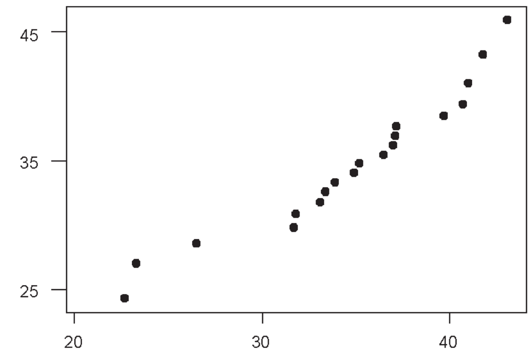


Figure 8

by multiplying the normal scores by s, the sample standard deviation, and adding $\bar{x}$ to this product. The reference set would be (Table 3):
and a plot of the data versus the reference set would produce Figure 8.

The plot of data from a normal distribution against their associated reference set should look like

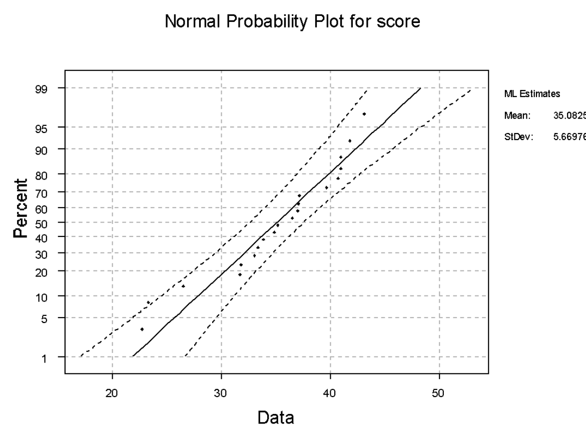


Figure 9

Table 3

36.1641	28.5273	35.4411	39.3986	36.9047	33.2603	43.2483
32.4880	40.9949	37.6770	26.9167	34.7239	34.0009	31.6656
24.1971	38.4994	30.7664	29.7472	45.9679	40.9949	

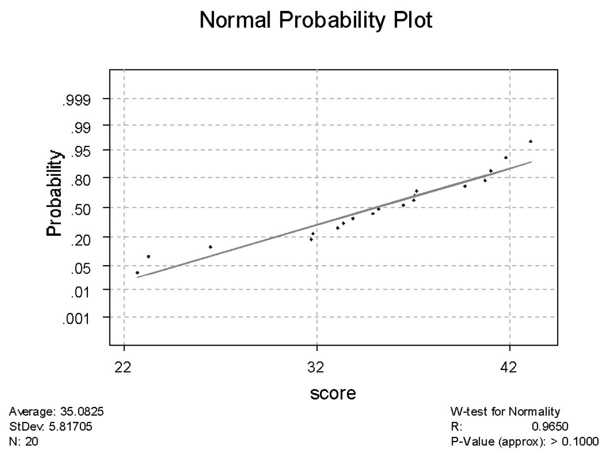


Figure 10

a straight line. This plot (without explicit mention of the reference set) is called a normal probability plot, and standard statistical packages produce this as a matter of course. MINITAB's normal probability plot of the data produces Figure 9.

Sometimes this plot is accompanied by a formal test of the hypothesis of normality. MINITAB's plot, accompanied by the Ryan-Joiner test statistic, is shown in Figure 10.

The plot, as well as the test, leads us to conclude that one can accept the hypothesis that the data came from a normal distribution.

The Central Limit Theorem has made unnecessary the understanding of the nature of the distribution from which the data were drawn when making inferences about the population mean based on the sample mean. All one needs to verify is that the observations are independent. So a good next step would be to teach nonparametric tests of independence. But, as one moves on to learning how to make inferences about other parameters, knowing the nature of the distribution from which the data were drawn is extremely critical. So, to answer the question posed in the title of this paper, we should next learn about normal probability plots, after which one can branch off into learning about inference about the population median, standard deviation, variance, or what-have-you.

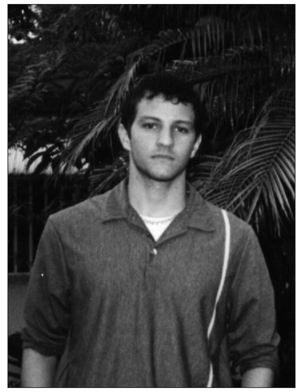
■ References

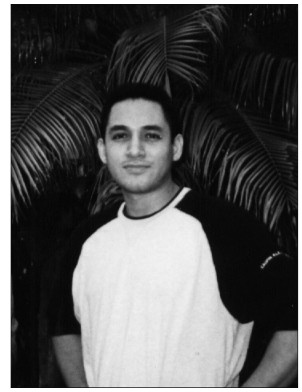
Madansky, A, (1988), *Prescriptions for Working Statisticians*, New York: Springer-Verlag.

Mosteller, F., Fienberg, S.E., and Rourke, R.E.K. (1983), *Beginning Statistics with Data Analysis*. Reading, MA: Addison-Wesley.

Student Project

Major League Soccer: Predicting Attendance for Future Expansion Teams





Diego De Rose Rafael Galarza

■ Introduction

As soccer fans, we are very interested in seeing Major League Soccer (MLS, the professional soccer league in the United States) become an accepted professional league in a country not known for its love of soccer. MLS began play in 1996, attempting to use the success of World Cup '94, held in the U.S., to create a league that would hopefully flourish in this country. The league opened with 10 teams mainly in major markets, with a few exceptions: Columbus, OH; Colorado (Denver); Tampa Bay; New York/New Jersey; Los Angeles; San Jose, CA; Dallas; Kansas City; New England (Boston); and Washington, D.C. In 1998, MLS expanded to 12 teams with the additions of franchises in Chicago and Miami. The league is expected to further expand into other cities in the near future. This is where we come in. Undoubtedly, the measure of how well a league does—and a team, for that matter—is through one statistic: attendance. "The first four MLS seasons drew a total of more than 11 million fans—an average of more than 15,000 fans per game, eclipsing the league founders' initial projections of 10-12,000 per game" (*www.mlsnet.com*). Nonetheless, some teams have done much better than others in this category, for varying reasons. MLS would like to expand into markets that will fill the stands as much as possible.

■ Project

The purpose of this project is to find an appropriate regression model relating average attendance for the 12 existing teams to factors that may influence attendance. Once the optimal regression model has been obtained, we will insert 16 promising cities into the regression to ultimately arrive at plausible expansion sites. This would give MLS an idea of which cities would generate good attendance figures if awarded a franchise. The majority of the 16 possible cities are major markets. But because teams in the smaller markets have been successful, some smaller markets will be analyzed as well.

Several assumptions have to be made before the project is initiated. First, some of the teams have attempted to boost attendance by attaching another event to the game, such as another game or a concert. We will assume a boost in attendance for one game would not drastically change the average attendance. Second, and most obvious, a team's win-loss record always affects attendance. But this information does not exist for possible expansion teams. And although this variable would help explain variation in attendance for current teams, our purpose is to predict attendance for potential expansion teams. Thus, only information that exists for future expansion teams can be studied. Finally, the average attendance was calculated from the 3.5 years the league has existed. Though time trends are impossible to measure in a short 3-year span, all extractable information was used.

■ Materials, Methods:

We attempted to predict how well a team fares in the stands using factors that logically relate to attendance. We brainstormed for tangible factors, realizing that there are other attendance-related variables such as player injuries and other local events that cannot be measured, and finally selected nine (Table 1). They are:

- Total Population (1997)
- No. of Professional Sports Teams in the four major sports

Diego De Rose and Rafael Galarza are graduates of Florida International University with Bachelor's Degrees of Science in Statistics.

- Unemployment Rate (1996)
- Consumer Price Index (1995)
- Third Quarter Cost-of-Living Index (1995)
- Mean Temperature (April–September, 1961–96)
- Total Precipitation (April–September, 1948–Present)
- Percent Hispanics (1996)
- Average Income (1997)

Many hours went into gathering these data from Internet resources. A few words need to be said about these data. First, it became difficult to calculate total populations for certain markets that have small cities in their outlying areas. For this, we came to the conclusion that any city within 40 miles of the team's actual market should be added into the measurement of population. We calculated that any distance over 40 miles meant a drive of more than one hour, which we figured as a boundary (Table 2). Secondly, Mean Temperature and Total Precipitation are calculated from April to September, when the league is in play. Although many weather factors exist, Mean Temperature and Total Precipitation were chosen because they reflect the most practical weather data, plus we feel they provide an adequate distinction between the weather patterns in the 12 cities.

After the data were collected, Minitab was used to perform least squares regression analysis. The first step was to run regressions on each of the nine individual factors. Incidentally, Total Population was the most valuable and Percent Hispanics was the least useful. Our initial belief was that cities with large Hispanic bases would fare better in the stands, but Miami's low average attendance is a prime example of the inaccuracy of our assumption. Our objective here was to weed out the less effective factors and begin to recognize the most effective. Overall, except for the low productiveness of Percent Hispanics as a factor, there were no surprises in this first run of analyses. A quick review of the individual plots showed no peculiar relationships. Although difficult to inspect with only 12 observations, some of the plots seemed relatively linear while others were highly scattered, which gave us an idea of potential factors.

A transformation of the predictors to the second order usually increases R^2 and the overall quality of a model because the squared terms take into account any possible curvature in the scatter plots. Thus, regression analysis also was run on all squares of the predictors. However, this proved ineffective, confirming our previous assumption that relationships were mostly linear. To get an early idea of any inherent multicollinearity, Pearson's correlation was run on all factors. Multiple regression was the ensuing logical step in acquiring a reliable model. Best Subsets Regression also was run just to get an idea of how many predictors would be necessary (Table 3). We decided that our model would have three predictors in addition to the constant. We came to this decision after observing a significant jump in R^2 (84.7% to 91.4%) from a 2- to a 3-predictor model, respectively, in Best Subsets. We wanted to have three or four models that were worthy of being chosen; these models were to be examined and one was to be selected.

■ Statistical Results

After analyzing many different models, we arrived at the one that we labeled the "best." Our first step in selecting the model was making sure all necessary assumptions were met. Pearson's correlation was run on all factors, and one pair showed a high correlation (Table 4): Total Population / No. of Teams. Right here, we realized that multicollinearity may be a problem in our model. We stuck with it for two reasons. First, the Variance Inflation Factors were not extremely high. VIF's are essentially the coefficients of determination of each independent variable with respect to all others. A large VIF means that the variance of the coefficient is high, which makes it unstable. A general rule of thumb says over 10 is high (*http://www.udel.edu/htr/Statistics/ Notes816/notes.html*). Second, multicollinearity is less of a problem if the aim of the regression is to predict a Y value rather than to find estimates of the coefficients.

Next, we checked whether the model violated any other multiple linear regression assumptions. A normal probability plot of residuals displayed that the data came from an approximately normal distribution (Figure 1). To test whether there was a problem with non-constant variance, we checked the plot of residuals versus fits for the model (Figure 2). No evidence of non-constant variance was visible. Hence, our model contained Total Population, No. of Teams and Mean Temperature (Table 5).

Thus, in our subjective opinion as researchers, our "best" model is:

$$\underset{(5178)}{Y = 28721} + \underset{(0.0002519)}{0.0013498\, X_1} - \underset{(387.8)}{972.4\, X_2} - \underset{(73.64)}{238.17\, X_3}$$

where X_1 is the Total Population, X_2 is the No. of Teams, and X_3 is the Mean Temperature.

With an extremely significant regression (F value 28.47), notably high R^2 and adjusted-R^2 for practical data (0.914 and 0.882, respectively), a healthy residual plot and somewhat small standardized residual values, this model will be

Table 3. Best Subsets Regression

Response is Avg Atte

Vars	R-Sq	Adj. R-Sq	C-p	s	Total Po	No. of T	Avg Inco	Unem-ploy	Mean Tem	Tot Prec	CPI	COL inde
1	63.9	60.3	12.0	2629.8	X							
1	37.9	31.7	26.5	3450.8							X	
1	37.7	31.5	26.6	3455.2					X			
2	84.7	81.3	2.5	1805.1	X				X			
2	80.2	75.8	5.0	2052.0	X	X						
2	74.7	69.1	8.0	2319.7	X					X		
3	91.4	88.2	0.8	1432.7	X	X			X			
3	86.7	81.7	3.4	1785.1	X			X	X			
3	86.6	81.6	3.4	1790.4	X	X	X					
4	93.0	89.0	1.9	1387.0	X	X		X	X			
4	92.3	87.9	2.3	1451.2	X	X			X			X
4	91.8	87.1	2.6	1500.5	X	X	X		X			
5	94.4	89.8	3.1	1332.7	X	X		X	X			X
5	93.4	87.9	3.7	1453.4	X	X		X	X		X	
5	93.3	87.8	3.7	1458.1	X	X	X	X	X			
6	94.5	87.9	5.0	1450.8	X	X	X	X	X			X
6	94.4	87.8	5.1	1459.7	X	X		X	X		X	X
6	94.4	87.8	5.1	1459.9	X	X		X	X	X		X
7	94.5	85.0	7.0	1618.2	X	X	X	X	X		X	X
7	94.5	85.0	7.0	1619.0	X	X	X	X	X	X		X
7	94.4	84.7	7.1	1631.9	X	X		X	X	X	X	X
8	94.6	80.2	9.0	1858.8	X	X	X	X	X	X	X	X

used to predict which of the 16 possible cities would produce the best average attendance.

■ Conclusion

Armed with a useful regression model, we inserted 16 deserving cities into our equation to see which would generate the best average attendance. The selection of these 16 markets was simple. Most are major markets that house millions of people; others are smaller markets that have successful franchises in at least one other sport. Las Vegas was considered only because we were curious as to how it would perform. Having been run through the regression and visually studied in case of any inaccuracies, four cities were selected as viable expansion sites (Table 6).

These cities are:		95% Prediction Interval
Philadelphia, PA	(15693)	(14774, 16611)
Detroit, MI	(15449)	(14086, 16811)
Portland, OR	(15384)	(13594, 17173)
San Diego, CA	(14553)	(13349, 15756)

As for our opinion on the results, Detroit and Philadelphia make sense because of their large populations. Portland and San Diego were a bit more surprising, but not too surprising considering their pleasant weather and modestly large populations. Thus, in conclusion, if MLS were to expand in the near future, these four markets would most likely have solid attendance figures.

Minitab, SAS and SPSS were used extensively throughout this project. All three were necessary tools in determining our final result. Finally, we would like to thank Dr. Sneh Gulati, Dr. Zhenmin Chen, Melina De Rose and Anthony Marquetti for assisting us with our project.

Table 4. Correlations (Pearson)

	Avg Atte	Total Po	no. of t
Total Po	0.800		
	0.002		
no. of t	0.483	0.861	
	0.112	0.000	
Mean Tem	-0.614	-0.211	-0.006
	0.034	0.510	0.986

Cell Contents: Correlation
P-Value

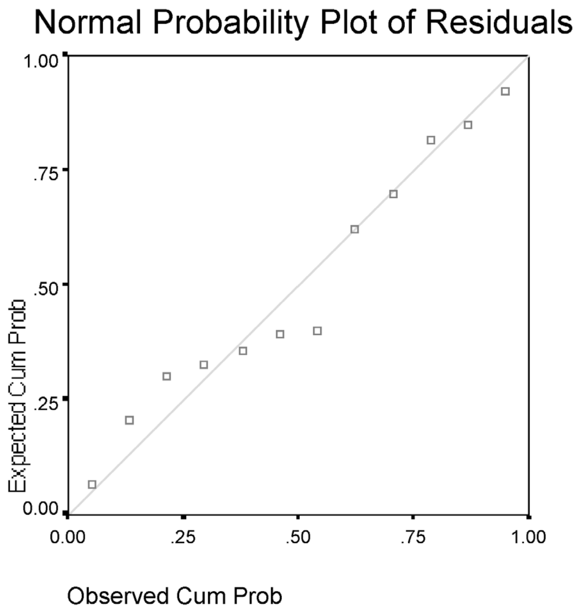


Figure 1

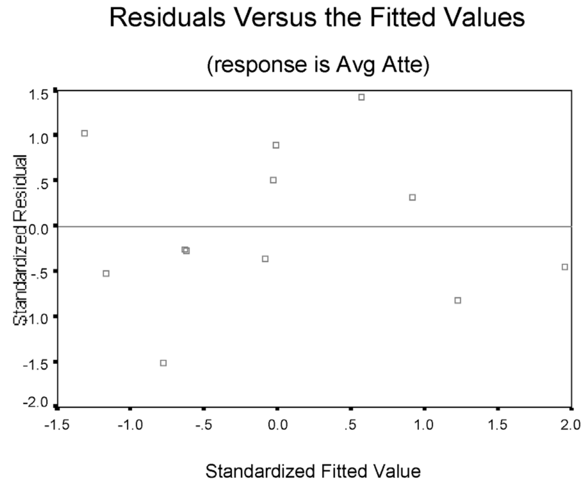


Figure 2

Table 5. Regression Analysis

The regression equation is Avg Atte = 28721 + 0.0013498 Total Po – 972.4 no. of teams – 238.17 Mean Temp

Predictor	Coef	StDev	T	P	VIF
Constant	28721	5178	5.55	0.001	
Total Po	0.0013498	0.0002519	5.36	0.001	4.621
no. of t	-972.4	387.8	-2.51	0.037	4.415
Mean Tem	-238.17	73.64	-3.23	0.012	1.196

S = 1433 R-Sq = 91.4% R-Sq(adj) = 88.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	175293376	58431125	28.47	0.000
Residual Error	8	16421311	2052664		
Total	11	191714687			

Source	DF	Seq SS
Total Po	1	122557007
no. of t	1	31262232
Mean Tem	1	21474136

Obs	Total Po	Avg Atte	Fit	StDev Fit	Residual	St Resid
1	1456440	15765	15024	895	741	0.66
2	4609414	16398	15111	495	1287	0.96
3	11780067	22340	22976	1062	-636	-0.66
4	3601914	9729	10481	859	-752	-0.66
5	1716818	9885	12059	563	-2174	-1.65
6	11148388	19273	18824	1094	449	0.49
7	4677781	12291	12652	689	-361	-0.29
8	1620453	14309	14836	783	-527	-0.44
9	2224973	11402	9922	797	1480	1.24
10	8506048	18896	20064	673	-1168	-0.92
11	2163511	12283	12672	1140	-389	-0.45
12	5826816	19486	17435	559	2051	1.55

Table 6. Predicted Attendance

	Total Population	No. of Teams	Mean Temp	Predicted Attendance	95% P.I. Lower Bound	95% P.I. Upper Bound
Philadelphia	5269580	4	68.2383	15693	14774	16611
Detroit	4468503	4	64.7217	15449	14086	16811
Portland	1789790	1	62.0583	15384	13594	17173
San Diego	2723711	2	66.76	14553	13349	15756
Pittsburgh	2359824	3	64.2083	13697	11992	15402
Houston	3846996	2	78.9833	13158	10978	15338
Baltimore	2475952	2	71.8333	13010	11837	14183
Indianapolis	1504451	2	66.97	12857	11453	14261
St. Louis	2559065	3	70.5133	12464	11299	13628
Atlanta	3634245	4	72.9583	12361	11140	13582
Las Vegas	1262427	0	79.3683	11522	8903	14141
Charlotte	1351675	2	72.0783	11434	10151	12717
Nashville	1136607	2	71.9617	11171	9839	12503
San Antonio	1506573	1	79.26	10905	8784	13026
Jacksonville	1028832	1	77.36	10712	8877	12547
Phoenix	2842030	4	84.8567	8457	6054	10860

Student Voices

Start an IV – STAT!

Amanda Elizabeth Brooks

Medical emergencies and traumatic injuries affect the lives of thousands of people daily. The lives of the ill and injured are directly affected by the interventions of medical professionals. The ability of these professionals to administer appropriate medications, whether lifesaving or pain-relieving, often depends upon the establishment of intravenous cannulation. Statistical methods can play an important role in assessing the efficiency with which Emergency Medical Services (EMS) personnel establish intravenous cannulation.

As a volunteer Emergency Medical Technician with Apex EMS in Apex, North Carolina, I have had the opportunity to observe the establishment of intravenous cannulation several times. I incorporated my fascination with EMS into the classroom by applying my observations in my high school AP Statistics class. A major portion of the final grade in this class depended on the development and appropriate analysis of a statistics project. For this assignment, I devised a statistics project titled *Success Rates of Intravenous Cannulation in Emergency Medical Services*.

For clarification, intravenous cannulation is the introduction of a needle surrounded by a catheter into a vein. Once the needle and catheter are in the vein, the needle is withdrawn. This leaves the catheter, a small flexible piece of plastic, in the vein. This process provides direct access to the blood stream for the delivery of fluid therapy and medications. It is more commonly called an "IV." The success rate of intravenous cannulation by EMS personnel is the proportion of attempts to start an IV that result in a free-flowing line in a vein.

I thought it would be intriguing to discover whether stress associated with life-threatening situations affects the ability of EMS personnel to successfully establish an IV. From my work-related experience, I suspected that intravenous cannulation success rates were lower under the stress of emergency situations. I wrote letters to Glenda Haynie, my statistics teacher at Southeast Raleigh High School, and Nicky Winstead, Apex EMS Administrator, requesting approval to carry out a project that would test this hypothesis. The project not only sufficed as a statistics grade for my course, but also provided quality control and education for Apex EMS.

This observational study was carried out at Apex EMS through the review of call records. Convenience sampling best defines the sampling method used; Apex EMS is just one EMS agency and cannot be used to make inferences about all EMS agencies. Counts of IV successes and failures were recorded, along with the number of attempts made to establish the IV. Then, each call was marked as an emergency or non-emergency call depending on whether lights and sirens were used during transport to the hospital. This data collection included call records from the months of January through August of 1999.

The data are summarized in Table 1. A chi-square test was performed to compare the success rates of intravenous cannulation in emergency and non-emergency situations. To my surprise, the comparison of proportions indicated that intravenous cannulation had higher success rates in emergency transport situations than in non-emergency situations. The most dramatic difference was found in the success of intravenous cannulation on the first attempt. Under emergency conditions, 94% of first attempts were successful; under non-emergency conditions, only 80% of first attempts were successful. A two-sample test for proportions supported a significantly higher

---

Amanda Brooks graduated from Southeast Raleigh High School in June 2000. She is a first-year student at the University of North Carolina at Chapel Hill. *She would like to enclose a special thanks to Jacquelin Dietz, statistics professor at North Carolina State University; Terri Griffin, Paramedic; Mike Fall, Paramedic; David Garrard, Paramedic; Shawn Mitchell, Paramedic; Greg Edwards, Paramedic; and Deborah Brooks, her mother.*

Figure 1. Apex EMS ambulance

success rate of intravenous cannulation on the first attempt in the emergency setting.

The research for this study began with the hypothesis that intravenous cannulation success rates were greater in non-emergency situations without the effects of stress and life-threatening patient conditions. Through careful research, tabulation, and testing, I discovered that the results were exactly the opposite of what I expected. My results show that intravenous cannulation success rates are higher in emergency situations. This finding reflects well upon the EMS personnel and the proficiency with which they perform under pressure.

Figure 2. IV bag and line hanging in the back of an ambulance, ready for use

This project was honored with a national award, winning first place in the American Statistical Association's 2000 American Statistics Project Competition. The research and statistical foundation needed for this project enhanced my overall understanding of mathematics and statistical models. In fact, my AP Statistics class was an experience that seemed to draw together several key concepts studied in high school, by enhancing such fundamental skills as data entry, organization, testing hypotheses, and problem solving. The construction of a solid foundation in AP Statistics lay in understanding and applying these concepts.

In summary, my study disproved my hypothesis that intravenous cannulation success rates were lower under the stress of emergency situations. Through tabulation of data and statistical testing the study proved that intravenous cannulation success rates were markedly higher in emergency situations. Ultimately, the study *Success Rates of Intravenous Cannulation in Emergency Medical Services* was a wonderful success, in spite of the unexpected outcome.

Figure 3. Paramedic Terri Griffin demonstrating how to establish an IV

Wake County Procedures

IV Access

INTERMEDIATE and PARAMEDIC

Procedure:

* Use peripheral IV sites (including the external jugular for adults as last option).
* Paramedics can use intraosseous access where threat to life exists for pediatric patients under 72 months old.
* Use the largest catheter bore necessary based upon patient's condition and size of veins
* Fluid and setup choice is preferably:
* Lactated ringers with a macrodrip (10 gtt/cc) for trauma or hypovolemia
* Normal saline with a macrodrip (10 gtt/cc) for medical conditions, and
* Normal Saline with a microdrip (60 gtt/cc) for medication infusions.
* Rates are preferably:
* Adult KVO: 60 cc/hr (1 gtt/ 6 sec for a macrodrip set)
* Pediatric KVO: 30 cc/hr (1 gtt/ 12 sec for a macrodrip set)
* If shock is present:
* Adult: 250 cc fluid boluses repeated as long as lungs are dry and BP < 90 Consider a second IV line.
* Pediatric: 20 cc/kg boluses repeated PRN for poor perfusion

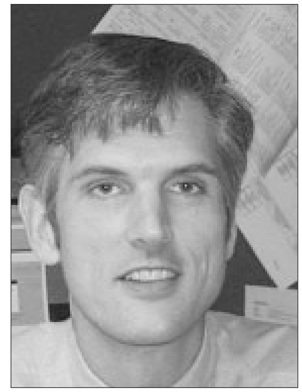
Figure 4. Wake County procedure for establishing intravenous cannulation

Table 1. Counts of Successes and Failures in First and Second Attempts at Intravenous Cannulation Under Emergency and Non-Emergency Conditions

	Emergency IV			Non-Emergency IV		
	Success	Failure	Total	Success	Failure	Total
1st attempt	65	4	69	253	64	317
2nd attempt	11	11	22	23	36	59

On the Job

A Day in the Life of a Statistical Manager at an Aerospace Company



Eric Fox

Having always liked mathematics, ending up with a career in statistics should have been no surprise to me. However, no one ever dreams of growing up and becoming a statistician. I started out majoring in computer science at Purdue University. After about one year, I discovered that I wasn't really interested in compilers and how a computer worked but rather in the mathematical algorithms used in programming. For this reason, I then switched my major to mathematics. Also, because of a microeconomics elective that I took, enjoyed, and did very well in, I simultaneously entered into the Economics Honors program. Although I found both to be very interesting, they just didn't "click" with what I ultimately wanted to do. My roommate at the time had just started in the Master's program in statistics, and I found myself looking through his books. Well, to make a long story short, I, too, decided to pursue a Master's Degree in Statistics from Purdue and graduated in 1987.

The summer after graduation, I took a position with Pratt & Whitney in West Palm Beach, FL - particularly appealing since I never did seem to like Indiana winters. Pratt & Whitney, a division of United Technologies, is a worldwide leader in the development and manufacture of jet and rocket engines. Pratt & Whitney has gas turbine engines in service with 370 airlines around the world. These engines power 60% of all western-built jet transports.

I worked as an engineering statistician for seven years before being promoted to the supervisor of the Engineering Statistics Group here in Florida. After being in that position for four years, I was promoted to my present position as Manager of Statistical Methods at Pratt & Whitney.

This position involves group management of statisticians in both Florida and Connecticut, development of strategies and policies for the use of statistical methods throughout Pratt & Whitney, and working on technical projects.

A typical day of mine is hard to define. I have some days where all that I do is group management. Other days may involve working on a single technical project. Still others involve a combination of the two. On a recent "typical" day, I arrived at my office around 8:00 AM. There were a couple of items on my desk that required immediate attention – a request from the boss, an expense report to check over and sign, and a request from a statistician to attend an upcoming conference. I also caught up on my e-mail.

At 9:00 AM, I met with two engineers from an engine program office who were working with a casting vendor to improve the casting yield for a turbine blade that they were producing for us. We decided that a designed experiment was the best approach to solve this problem. However, due to the complexities of the process, it was decided that we would need to travel to Michigan to walk through the process and to interface with their engineers as well. But based on the information received so far, I indicated that I'd put together a couple of preliminary designs (probably fractional factorials) that we could build upon which would be used to stimulate discussions for the final designs. Based on everyone's schedule, we decided to meet with the supplier in Michigan in two weeks.

One hour later, I had a telephone conversation with one of the statisticians from Connecticut regarding modifications to some sections of our 4-day Design of Experiments (DOX) Workshop that we teach four times per year both in Florida and in Connecticut. We decided to add some

Eric P. Fox is Manager, Statistical Methods, Pratt & Whitney

information to the introduction summarizing more recent DOX successes that we've had rather than the older cases that were presently in the workbook. In addition, we determined that we needed to give more detailed coverage of why DOX was a better problem-solving tool than some standard methods such as one-factor-at-a-time experimentation. We decided to add a two variable example, which would show how a designed experiment would find an optimum but where the one-factor-at-a-time approach would find a sub-optimum point.

At 10:30 AM, I had a regularly scheduled weekly status meeting with one of the statisticians in the Florida group. This is time that we use to update each other on the status of our jobs and workload and when we address any technical or personal issues that may be present. Of course, everyone is free to stop by anytime when I am not busy to discuss problems, but our weekly status meeting is a time that we put on our calendars each week. Today, we discussed a risk analysis that this statistician was conducting to predict the number of future cracked parts that we would expect to see in different maintenance approaches, with and without a proposed inspection of the part. We determined if the results made sense and discussed whether all of the assumptions of the analysis would stand up to scrutiny.

Forty-five minutes later, I had a second regularly scheduled weekly status meeting with another statistician in the Florida group. We discussed a reliability demonstration test plan that the statistician was working on and strategies to possibly reduce the number of tests that were required. As part of this discussion, we considered reliability growth models such as those found in MIL-HDBK-189. In fact, the US military has used demonstration test plans since the late 1950's to insure the mission reliability of its aircraft in a wide variety of scenarios. There exist military standards, such as MIL-STD-817, that are completely devoted to the conduct of reliability demonstration tests and the analysis of the resultant data. In the early days, these test plans were based on times to failure following an exponential failure model and the number of failures in a set test time following the Poisson law. Today such revised standards include a wide spectrum of possible models in the Weibull, lognormal, gamma, extreme-value, etc. Methods for analyzing these data include advances in Bayesian inference and other methods as well. In addition, we discussed a probabilistic design analysis that the statistician was working on for a rocket customer. Probabilistic design involves a careful examination of the main possible causes of component failure and the subsequent engineering design methods used to eliminate the occurrence or diminish its prevalence. This involves the use of a computer package that the customer has developed to predict life and incorporates uncertainties to the model input variables so that a distribution of life can be derived.

At noon, my boss called a staff meeting. I never do like to work through lunch, but he's always very good at providing pizza, salad, and soft drinks. We discussed numerous issues at the meeting including a merit budget for the upcoming year, problems with some employees getting their charges for the previous day into the timekeeping system on time, and coordination issues with some of our counterparts in Connecticut.

At 1:00 PM, I walked back over to my office and stopped by to talk to the statisticians whom I hadn't yet said "hi" to that day. It's amazing how often we all get busy and may not see each other every day. About half of the group was at their desks at that time. It seems like 1:00 PM is a popular time to call meetings. One statistician was having problems getting a matrix decomposition algorithm to work within a correlated random variable simulation. We worked on it for 30 minutes until we finally found the problem.

After getting back to my office, I used 20 minutes before a scheduled 2:00 PM teleconference to read 20 e-mail messages that somehow managed to arrive since I last checked earlier in the morning. Two were for new jobs that were being requested — one was a logistic regression analysis and the other was for a censored lifetime data problem. I forwarded the notes to two separate statisticians to see if they could work the jobs into their schedules. Another note requested budgeting requirements for the Statistical Methods Group for the year 2000 for a specific engine program. This would take a little bit of time for me to work through so I filed it for retrieval at a later time.

My teleconference for the Society of Automotive Engineers Probabilistic Methods Committee began on schedule. I am the vice-chairman of this committee of 130 members, and the leadership meets via teleconference monthly between our semi-annual meetings. The committee's purpose is to develop industry standards for the use and implementation of probabilistic methods primarily within aerospace and automotive industries, government, and academia. This particular meeting discussed the status on several Aerospace Information Reports that were being worked by many of the project teams. We also discussed an upcoming workshop titled "Practical Implementation of Probabilistic Methods" that our committee was sponsoring.

At 3:00 PM, I took some time to work on putting together an agenda for an upcoming Statistical Methods Offsite Meeting. This semi-annual meeting allows all Florida-based and Connecticut-based statisticians to get together for two days of technical interchange. At our next offsite, we will be hearing journal article review presentations, learning statistical methodologies from overviews of workshops that have been recently attended, and brainstorming ideas on ways that we can improve customer satisfaction. In addition, we will have the opportunity to tour our engine hardware museum.

At 3:30 PM, I had some time to work on a memo for a bootstrapping algorithm that I recently worked on with an engineer from our Structures Technology area. It involved the life prediction of composite materials. Having only 30 minutes to work on it would not allow me to finish the memo, but I was able to get a good start on it.

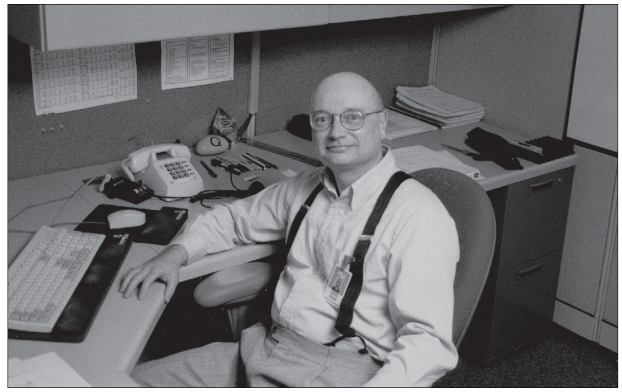
I stopped by my boss's office around 4:00 PM to remind him that I would be up in Connecticut for the next two days interfacing with the statisticians there. In addition he wanted to discuss a couple of items including one personnel issue and a question to me about my recruiting plans in Connecticut and Florida for the upcoming year.

At 4:30 PM, I began to pack my briefcase and included information that I would need for my meeting with the statisticians in Connecticut for the next two days. In addition, I threw in some notes that I made to myself about a paper that I'm working on about computational methods for probabilistic design for an upcoming conference six months away. Although the final paper isn't due for four more months, I like to keep chipping away at it since I know that the deadline will creep up quickly.

I left for home to pack for my trip to Connecticut at 4:45 PM. I wish that I could be one of those individuals who can plan ahead a bit better, but alas I always seem to be finishing things up at the last minute. I arrived at home at 5:15 PM, got my belongings packed, grabbed a quick bite to eat, and headed to the West Palm Beach airport just in time to catch my 7:30 PM flight to Hartford. On average, I travel about four to six days per month, which is a combination of company-related business, meeting with group members in Connecticut, and attendance at society meetings, technical workshops, and conferences.

All in all, the job as Manager of Statistical Methods for Pratt & Whitney is both exciting and challenging. Although I don't do as much technical work as I used to, it is equally challenging to determine how statistical methods will be used at a large company. But without a doubt, one thing stands out — my entire statistics degree is used all of the time at Pratt & Whitney due to the incredible variety of work.

A Day in the Life of a Statistician at Intel



David Drain

Intel was founded in 1968 to build semiconductor memory products; today, Intel supplies the computing industry with the chips, boards, systems and software that are the "ingredients" of computer architecture. Intel employs about 65,000 employees worldwide and had 1998 net revenue of 26.3 billion dollars. The Chandler site houses fabrication (fab), sort, assembly, and test manufacturing and a variety of development and support functions.

I joined Intel Corporation in Arizona in 1987, and I can state with certainty that it has never been a dull experience. I had worked in the semiconductor industry before—both as an internal consultant and as a simulation programmer—but this was the first time I had worked in a manufacturing environment. My Master's degree in Applied Statistics and prior work experience gave me most of the technical background I needed except for a notable gap in the area of Statistical Process Control (SPC), so I spent much of my first year at Intel learning about this field. The manufacturing processes were also very novel to me—solid-state diffusion, plasma etch, ion implantation, and micro-lithography, to name a few—so I worked closely with engineers owning those processes to better understand the problems we were trying to solve. I spent over six years in my first factory position, and most of my consulting related to SPC and Design of Experiments (DOE). DOE was routinely used to improve processes and evaluate proposed process changes. As semiconductor technology evolved, measurement capability became a more important issue and standards were devised for measurement capability analysis considerably more complex than the standard gage repeatability and reliability studies.

David Drain has been a Senior Statistician at Intel Corporation in Chandler, Arizona for the past twelve years. His main focus has been Statistical Process Control and Experimental Design applications in engineering environments. He has also written two books to support training engineers in the application of these tools. Prior to his employment at Intel, Mr. Drain worked as a statistical consultant, a simulation programmer, and a program development manager for a statistical software company. Mr. Drain has a Master of Science degree in applied statistics (1980) from Bowling Green State University in Bowling Green, Ohio.

My next Intel position was on the startup crew for a new factory that started production in 1995 and it was gratifying to see that, although the tools and technologies were very different from what I had seen before, the statistical tools I had to apply were nearly the same. The biggest difference of this new factory is that it is part of a virtual factory—a set of factories with the same tools and processes that work "virtually" as one large factory. Matching equipment inputs and outputs across this virtual factory occupied much of our time—both before and after production began. I recognized that one of the huge advantages of a virtual factory system is that any new problem need only be solved in one factory, and the solution can be copied to other factories immediately because they have the same manufacturing processes. Measurement capability continued to be an area where increasing demands and competition forced us to innovate and become more rigorous: we also developed methods to evaluate the state of equipment and process matching across factories and to accelerate equipment qualification during factory start-up.

My experiences in the factory motivated me to summarize the statistical knowledge that I believe engineers must have in order to be successful in a modern semiconductor production facility, and in late 1996 I completed two textbooks that I had started in 1992: *Statistical Methods for Industrial Process Control*, and *Handbook of Experimental Methods for Process Improv*ement (both published by Chapman & Hall). I have used these books in teaching both within Intel and elsewhere.

A little over a year ago I transferred into Intel's Logistics Group as their first statistician. This is a service organization and I find myself applying many new statistical methods which better fit a new and challenging set of problems. Categorical data analysis and industrial engineering applications such as time studies are common tasks for me now. Intel Logistics operations worldwide are supported by this group, so I find myself working with people all over the world. This brings a challenging new multi-cultural aspect to my job, and I have had to work hard to understand these differences to continue to work effectively. A typical day for me now may start at six in the morning and end around six at night—but I am able to flex my time and do some work from home so I still average a normal work week. Logistics has embraced the virtual factory concept described above down to the organizational practices and types of routine meetings. Once a week the world-wide Logistics Quality team meets; on the agenda are topics like discussion of recent audits in Malaysia, the best known methods for protecting shipments from water damage during hurricane season, or details on the costs of mis-delivery.

Another important aspect of my current job is to participate on problem-solving teams; one current team is working to prevent shipments from being damaged in transit. I am not an expert on warehouse and transportation details; my role is to facilitate the problem-solving process and help the team use analytic tools (such as fault tree analysis) to understand their problem and select solutions that will be effective in preventing damage. I often find that I am asked to do some training for these teams with no prior warning, so I've had to become adept at explaining complicated concepts to non-technical audiences using examples and experiences from their own lives—stories involving cars or dogs seem to work especially well.

I still do plenty of typical statistical consulting—helping to design work timing studies, or a DOE to evaluate layout changes, for example. This is one of the most satisfying parts of my job because it allows me to use statistics as a lever—with a little bit of data we can make a much better decision or save much more money. Measurement and measurement capability have a completely different meaning here in logistics than they did in the fab. This is a service environment, so I have had to learn how to measure things like customer satisfaction and service error rates; in the fab, measurement was technically more difficult, but the measurements were the well-defined continuous variables that classical measurement capability studies were designed to evaluate. I have helped individuals and groups understand how we measure errors and why we measure site performance with error rates rather than raw error counts,and I've written some general advice on measurement practices in our part of the company. I am also the Statistician on our Logistics Change Control Board—a body that reviews changes affecting Logistics operations around the world, and I am a member of the Escape Response Team that makes sure corrective actions are in place for especially serious incidents. We don't have a programming staff—as Intel employees we're expected to be familiar with common software tools—so I also write software and produce some routine reports.

I do lots of teaching—for people in logistics, for other Intel audiences, and for our vendors. One of my current projects is to design and develop a Statistical curriculum for logistics employees; I expect this will evolve into a more general curriculum for employees in other service-oriented parts of the company. I enjoy teaching so much that I also teach one semester a year at Arizona State University.

I have enjoyed working at Intel—I love helping people solve problems, and I think that the work I do here makes life better for people in our country and around the world. I especially enjoy the company of my fellow statisticians. We don't have a central statistics group—everyone works for the operation that they support—but we have a very active informal network that can provide technical expertise in areas where I find myself lacking. Often on a weekend I find myself out hiking, hunting, or maybe just planting a tree with another Statistician, and in spite of our resolve to do otherwise, we usually end up talking about Statistics. Our one formal activity is an annual Statisticians' Summit meeting—the last one was in Santa Clara and Douglas Montgomery (a professor at ASU and a prolific author) came in to talk about many different facets of DOE. Intel statisticians from all over the world also give presentations at these conferences, so they are a great opportunity to learn from one another.

People often ask me what it takes to be successful as a statistician in industry, so I'll tell you here the three qualities I look for:

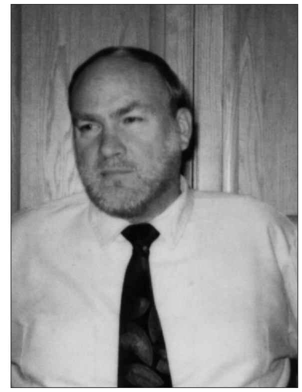
Communication skills, both written and oral, must be excellent, and these skills must include being able to restate statistical concepts in non-statistical terms for a variety of audiences.

A successful statistician must put the needs of the business first, and use statistics as a tool to meet those needs. Some new graduates seem compelled to apply every statistical tool they learned in school—whether the business requires it or not.

Technical skills must be sufficient: SPC and DOE at a minimum, and the more the better. Technical background outside of statistics—in chemistry, physics, and engineering for example—is especially valuable because it facilitates

continued on page 25

A Day in the Life of a Medicaid Fraud Statistician



Terry Allen

■ The Bureau of Medicaid Fraud

I am a statistician with the Utah Bureau of Medicaid Fraud, the agency charged with policing fraud and abuse in the $800 million-per-year state Medicaid system. In contrast to the Medicare program for which all elderly people qualify, Medicaid eligibility is based on the recipients' income and assets, focusing on the poor of all ages. The range of services paid by Medicaid includes visits to physicians, dental care, hospital and nursing home care, prescription drugs and mental health services.

Physicians and other Medicaid providers send their claims directly to the state Medicaid administrator to be analyzed by computer for completeness. Very little other review occurs. It is assumed that the services were actually provided, that they were necessary, and that they were honestly described on the claims. Claims are then paid, again by computer. Only in rare instances do human eyes review a claim to see if it appears honest and accurate.

Because of the difficulty in checking all billings in the Medicaid system, some providers embellish claims to supplement their incomes. Although Utah is a low volume Medicaid State it is estimated that as much as $80 million may be stolen annually by dishonest Utah providers.

Several years ago I was hired to create a statistical section for the Bureau of Medicaid Fraud to accomplish two goals. One goal was to organize the data into databases from which statistical summaries could be easily obtained. My second charge was to statistically analyze billing data to target suspicious patterns and to determine which providers warranted further investigation.

Today, both of these assignments are in place and ongoing. Investigators receive statistical summaries in a matter of minutes rather than weeks and statistical techniques have targeted more Medicaid fraud than our investigators have time to examine.

■ A Day in the Life

When I arrive at my office, the chief investigator, Jeff, is waiting to discuss a graph that I created for him some time ago (Figure 1). The bureau is preparing to file charges against a family practice doctor who works exclusively with nursing home patients. Jeff and I discuss the meaning of the graph and the impact it will have on the provider. We've found that well-executed graphics sometimes can help to avoid the time and expense of a trial. By summarizing the data into a clear picture, a number of graphs have convinced doctors and their attorneys to settle their cases out of court. The power of statistical graphics was emphasized for us in one situation when our investigators confronted a physician with his fraudulent billing pattern displayed in an easy-to-understand visual. The doctor became so upset that he started to hyperventilate and fainted. Our investigators found themselves in the peculiar position of having to revive the doctor in his own office before he could discuss his upcoded billings. After Jeff leaves, I begin to work on a new idea for targeting physicians who bill for surgical procedures when non-surgical treatments were actually used. This idea came from an investigator who noticed that a doctor had billed for a surgical circumcision when the patient record indicated that a clamp was used. Medicaid reimburses $100

Terry Allen is the Research Analyst for the Utah Bureau of Medicaid Fraud. He has an M.S. degree in Statistics from Brigham Young University and a Ph.D. in Mathematical Sociology from the University of Utah. He is an adjunct professor of sociology at the University of Utah, where he teaches graduate and undergraduate statistics courses. He also teaches classes in research methods, criminology, and deviant behavior. Dr. Allen has twice been elected president of the Utah Chapter of the American Statistical Association. He has published a number of papers on social issues and two on Medicaid fraud.

for a surgical circumcision and $32 for a clamp procedure—a tidy profit for doctors who upcode their billings! I begin my "statistical" detective work by looking for doctors who bill for a high percent of surgical circumcisions. I am interrupted by Michelle, one of our medical investigators, who needs a random sample of a doctor's patients. She is preparing a subpoena for the doctor's patient records. We want to know how widespread a particular illegal billing is. Since active cases take priority over statistical targeting, I locate the doctor's billing records and generate a statistical random sample of patients. Our investigator will use this list to request records from the doctor. She will then bring these sample files back to our office for auditing. Later we will use the overpayment found in the sample to estimate the total amount that Medicaid has overpaid the doctor. Back to the targeting. I begin by creating a database of doctors who have billed Medicaid for circumcisions in the past eighteen months. Next, I want to calculate the percent of surgical procedures for each physician. But first, I answer the telephone. The call is from the Department of Public Safety. They are requesting an item analysis on their Sergeants' and Lieutenants' Exam to determine which questions are good discriminators and which should be reworded or eliminated. As the only statistician in the Utah Department of Public Safety, I get asked to do this kind of project from time to time. It gives me an opportunity to promote the value of statistics beyond the Bureau of Medicaid Fraud. I tell them to send over the raw data. After lunch I return to the office and meet with Steve, another medical investigator. He has received a complaint that some medical supply houses are billing Medicaid for expensive wheelchairs that have not been supplied to patients. In some cases, the wheelchairs were not needed. In other cases, less expensive wheelchairs were substituted. I generate a list of medical suppliers who have billed for wheelchairs costing more than $17,000 each. Back to targeting once again. I calculate a Z-score on the percent of surgical circumcisions for each physician and generate a list of those who are more than two standard deviations above the mean. Twelve doctors appear on this list. I am interrupted this time by Denis, the Assistant Attorney General, who prosecutes our cases. He is going to court and needs a graph showing how the defendant, a pediatrician, has a very different billing pattern than his peers. I targeted this case some months ago, and it has finally worked its way through the system. I prepare the court graphic and begin to generate a circumcision bar graph to illustrate the extremely high percent of surgical procedures billed by the twelve outlier physicians (or out-and-out liars, as some of our investigators affectionately call them). The graph will be used by Jeff to decide which doctors' records should be examined by our investigators. I notice that the building is very quiet. Everyone else has left for the day. I spend half an hour running the back stairs for exercise, drop the new graph on Jeff's desk, set the alarm system, and head for home.

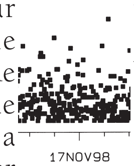


Figure 1. The graph shows the

A Day in the Life of a Pharmaceutical Industry Statistician

Stephen Gulyas

Being a biostatistician at a pharmaceutical company is a unique professional opportunity that is very gratifying. Knowing that you are working directly to improve people's lives is a benefit that nearly everyone I have met within the industry cherishes. Thanks in part to the Kefauver-Harris amendments passed by Congress in 1962, which mandated the use of clinical trials to demonstrate efficacy and safety of investigational drugs, statisticians are an integral part of the analysis of clinical trials and in the development of each drug candidate. If the results of an analysis show that a new drug is promising, the company may proceed with further development: a very costly, time-consuming, and risky investment, yet one which has the potential of helping millions of people worldwide. However, if the data do not support such conclusions, we can provide the information critical for management to stop development in its tracks and presumably spare patients from an ineffective treatment.

So you can understand some of the day-to-day workings of the industry, I will first provide a brief summary of the drug development process. Each year, tens of thousands of drug candidates are created and assessed in our discovery laboratories via procedures such as high speed analoging and high throughput screening. Biological mechanisms are studied and then novel chemical structures are constructed with some pharmaceutical purpose, also known as an "indication", in mind. An immense library is continually amassed from all of these candidates, and appropriate patents are submitted to the government to establish proprietary usage. This library, and consequently, a pharmaceutical company's "pipeline" of drug candidates, is the key to its long-term success. Once the structure is created, many parallel activities occur. Drug candidates proceed through many stabilization trials, dosage formulations, and studies in animals. If effective and safe, an IND (Investigational New Drug) application is submitted to the Food and Drug Administration (FDA) to request that the drug be tested in the human body. If the IND is accepted, the clinical development process formally begins.

Once the compounds reach humans, there are four well-accepted stages or phases of development. In Phase I, small studies (often 30 subjects or less) are generally run on healthy, normal volunteers. Properties of the drug such as the maximum tolerably safe dose and the rate and amount at which the drug is absorbed and eliminated from the body are ascertained. These studies often do not take very long to complete, and most finish in well under six months. Also, these studies generally use fewer patient recruiting centers than the other phases. Phase II uses larger studies (dozens to a few hundred subjects) to determine the proper dose and to verify proof-of-concept for efficacy on the target disease or condition. Efficacy refers to the drug's ability to combat the disease or condition of interest. These are often parallel-arm designs that are double-blinded and may take a couple of years to finish. Phase III studies are even larger trials (several

---

Dr. Stephen Gulyas has been a Coordinator in the Biometrics Department at Pfizer Central Research in Groton, CT since 1997. He received his M.S. and Ph.D. degrees in Statistics from Cornell University in 1995 and 1998, respectively, as well as a B.S. in Applied Mathematics from Carnegie Mellon University in 1992. Steve is an active member in the ASA, serving on the program committee and as webmaster for the Connecticut Chapter, and sitting on the Council of Chapters World Wide Web committee. In his free time, he plays and follows sports avidly, participates in musical activities, and enjoys travelling.

hundreds to thousands) that again seek to confirm efficacy and safety, to include a wider population representation by using many centers, and also to provide information about the incidence of possible adverse events. Another goal is to focus on various demographic subsets (e.g., older subjects, different races) for both efficacy as well as safety considerations. Studies that attempt to show equivalence or superiority to already-existing marketed drugs, to measure quality-of-life considerations, or to assess cost savings of various therapies through health economics may be investigated in either Phase II or Phase III.

If the drug passes all of these hurdles successfully, during the end of phase III, the drug is submitted to the FDA (in the U.S.) for regulatory approval. This is called a New Drug Application, or NDA. If all goes well, the drug is approved for the indications sought, a "label" describing all aspects of the drug is agreed upon, and finally the drug can be prescribed. Subsequently, phase IV studies may be run for post-marketing surveillance. Once a drug is marketed, a new indication for the drug may be sought; this might include another population not covered in the original label. Also, although often not part of a separate clinical trial, pharmaceutical companies are responsible for tracking unsolicited reports of spontaneous adverse events reported to the company by individuals. All in all, drug development is quite an intense process with literally thousands of scientific, logistic, and business pitfalls: approximately one of 100,000 compounds will survive to market. The total time from discovery to drug approval may range anywhere between seven to over twenty years. Figure 1 summarizes the entire process.

I am a biostatistician primarily responsible for proposing, negotiating, and conducting designs and analyses agreeable to the drug project team in phase II and phase III clinical trials. My therapeutic areas of expertise include oncology, anti-infective, and cardiovascular compounds. Different groups of statisticians support special phase I studies, nonclinical studies, and pure research activities, or may have roles within management. In my world, the following paragraphs chronicle a typical day I might experience:

8:07 a.m.: I arrive into my office and plug in my docking station to check email. Email is used extensively to communicate with internal and external contacts, and I receive about 40 messages per day. Fortunately, there isn't anything urgent in my inbox today, so it appears I will not have to rearrange my schedule. Before my 9:00 meeting, I call our help desk to resolve a problem about my Palm Pilot, answer voice mail, respond to a few emails, and look over the materials I will need for the day.

9:00 a.m.: Our weekly team meeting for a cardiovascular compound starts. This team is comprised of members from Clinical Research, Regulatory, Biometrics, Outcomes Research, Medical Writing, Pharmacy, and other departments. The members from the clinical team include a lead clinician (who has a M.D. degree) for the entire project, and one to three other clinicians who each has responsibility over individual study protocol(s). A protocol is a set of rules that describe how a clinical study should be run. There are five CRAs (clinical research associates) present today. They monitor the individual clinics and/or physicians that we contract with to recruit subjects and implement our studies, and they ensure that the prescribed rules in the protocol are being followed. The regulatory group interacts with the FDA or appropriate foreign regulatory body to report changes in the study, keeps the agency apprised of the entire clinical development program, and also considers strategic issues. Outcomes Researchers grapple with issues such as quality-of-life, patient satisfaction with the drug, and health economics. Medical Writers author final study reports, summarize findings across protocols for the entire drug project, and submit articles for publication. A member of the pharmacy is available to discuss issues in drug formulation and drug supply. The members from Biometrics include a team leader who serves in a "quarterbacking" role, statisticians, statistical programmers, and data managers who specialize in the archiving of efficacy or safety information.

The most relevant item for me on the agenda today is to report on the progress of revising a statistical analysis plan (SAP) for this protocol based on the team's earlier comments. Although a well-written protocol will itself contain a fairly detailed statistical methods section, the SAP is a document that spells out the full statistical details to be carried out in the final analysis. It describes which analyses are being planned, which variables are of interest, what to do about missing data, what the definitions of necessary algorithms are, etc. The SAP provides proof of ethical conduct around analyses since they are specified prior to seeing the data. Thus, the analyses cannot be construed by others as data-dredging, done only to salvage the life of an undeserving drug. When the method of analysis is nontrivial or not prescribed by a regulatory agency, a well-conceived, negotiated, and signed-off SAP is an appropriate vehicle for communicating analysis intentions and helps to maintain the integrity of the trial.

10:04 a.m.: I return from this meeting and the voice mail light on my phone is lit again. I check the messages and respond to a few urgent items. There is a message from a CRA who would like my input in designing a Case Report Form (CRF): the sheet used by a patient or investigator to record the appropriate measurements taken while participating in the study. A second message is a meeting request from a statistical programmer who wants me to explain in more detail sections of the SAP for a certain study. Another message is from an internal public relations member asking me to provide details for an in-house talk statistical I am organizing that will be given by an external speaker.

Finally, I have the remainder of the hour to tackle some technical statistical issues. I return a call from a clinician at a Pfizer office in Sandwich, England. We discuss the statistical ramifications of having proposed a two- versus three-parallel arm design. Statisticians perform lots of consulting when first writing a protocol to ensure that the questions our physicians want to have answered will be cleanly addressed in the design. This is a major part of how we contribute to the planning of studies. Also, working with, and travelling to meet, international colleagues is not uncommon. Since the U.S., Canada, European Union, Japan, and other blocks of countries all have individual regulatory agencies that handle drug submissions, it is critical that pharmaceutical companies have a major presence worldwide. Studies may also be run anywhere in the world to help speed recruitment, highlight demographic distinctions, and reflect differences in medical practice.

10:59 a.m.: Next on the agenda today is a statistical recruitment meeting. This team works bi-weekly to help our statistics group manage interview candidates, share our experiences with universities, improve our interviewing process, and work with the Employee Resources Department on new initiatives. Lately, we have been spending some time preparing for our activities at the 1999 Joint Statistical Meetings in Baltimore. We are participating in the placement service as well holding a separate open house. The open house is intended to allow us to share some of Pfizer's corporate vision, speak candidly about our experiences at Pfizer, spark discussion with professional colleagues from academia, regulatory bodies, and industry, and talk about potential future positions such as summer interns or permanent employees.

12:19 p.m.: It's time for lunch in our first-rate cafeteria. We have quite a selection of items from which to choose: stir fry to order, hot and cold deli sandwiches for take out, a full grill, brick oven pizza, and gourmet desserts. With all due respect to my alma mater's (Cornell) fabulous desserts, it sure beats any school cafeteria food I've ever had.

1:00 p.m.: I'm off to a technical session for professional development. This is a roundtable discussion that meets bi-weekly where most (about a dozen) of the people in the room are statisticians with less than 5 years of industry experience. However, there are a couple of statisticians with at least 12 years of experience that are there mainly to guide the discussion. A discussant prepares a few slides, and we invariably interrupt the presentation and talk about the topic in a stream-of-consciousness manner. The goal is not necessarily to arrive at any one conclusion but to hear the common arguments for or against a particular issue. In the past, we have talked about the phases of drug development, analysis of subset populations, data safety monitoring boards, interim analyses, missing data, etc. Personally, I find this one of the most useful forums for exchange of statistical issues that we have here at Pfizer.

1:57 p.m.: Now it's off to a biometrics team meeting about an oncology compound. This meeting consists of a biometrics team leader, three statistical programmers, a data manager, one safety data manager, and two people from Biometrics Operations Support. We have recently completed an analysis on two of the protocols we are running. Subsequently, we are dealing with the loose ends from that analysis and verifying that all the results have been properly reported. A third protocol must be further modified to allow for additional analyses akin to the ones recently performed on the other protocols, so the technical issues that will arise are discussed.

2:58 p.m.: I stop to pick up my U.S. mail for the day. I find the latest copy of *The American Statistician*, some junk mail, some Pfizer information about Y2K preparations, a memo about a change in the 401(k) plan, information about the Pfizer Employee Health & Fitness Center, and a Cornell alumni magazine. This will be some light reading for when I get home tonight.

3:11 p.m.: I have some more open time. I read up on a couple of articles that describe some of the biological mechanisms behind my assigned cardiovascular compound. Understanding the biology is as important for us as it is for the clinicians to understand the statistics we are using to present the results. While occasionally addressing an intermittent phone call, I put some more thought into the parallel arm design that I discussed earlier and derive contrasts for some additional questions the clinician has. This also sparks some questions about the data from a previous study I have analyzed, so I go back and create more graphs and run some logistic regressions in SAS, and record my findings. In reading some more email messages, I find some items that must be uploaded to the Connecticut Chapter of the ASA website that I maintain, and call a few of the other officers to discuss some questions I have. Participation in professional societies is stressed within the industry, as is evident by the fact that the Biopharmaceutical Section is the largest of the ASA sections. Publication in both medical and statistical journals is also valued.

5:47 p.m.: I boot down the computer and pack up for the evening. Although some times are busier than others, especially when a compound is undergoing an NDA submission, on the whole, pharmaceutical companies recognize the value of their people. They encourage employees to find the right balance between work and home lives.

Working in the pharmaceutical industry offers many challenges and opportunities for statisticians. Core research, applied methodology, and management leadership are all common areas where statisticians reside. Although the work may differ, nearly all positions require excellent consultation, collaboration, and communication skills. I am proud to be a biostatistician in this industry and look forward to being a key player in future successful drug submissions that will directly result in improving people's lives.

continued from page 19

communication with our client engineers and helps the statistician understand the context of the problem better.

People also ask me how to get a job in industry, so here are some suggestions: Take advantage of the many internships offered by industry—I had two interns last summer and I know they acquired some valuable skills during that time. Attend the ASA meetings, and if you are near graduation, sign up for the job fair there. Lastly, make sure your graduate school has a web-site that shows potential employers who is available for internships and full-time employment, and how to contact those people. See *http://www.k-state.edu/stats/gradstudents.html* for a great example of an effective web-site.

Outlier...s

Allan Rossman

I have chosen "rankings" as the theme of this issue's column. Rankings and "top ten" lists have become a favorite device for columnists and talk show hosts alike, and the coming of the year 2000 brought a plethora of lists of the "top ___" of the preceding century and even millennium. I will confine my attention to rankings that are at least marginally relevant to students of statistics, and I guarantee that this will be one of the four best "Outlier...s" columns that I have produced thus far. [**Assignment 1**: How many columns have I produced thus far? The answer appears at the end.]

■ College Rankings

Among the most important sets of rankings to students (not to mention their parents and college administrators) are college rankings. Perhaps the best-known and most influential are those produced by *U.S. News and World Report*, available on the Web at *http://www.usnews.com/usnews/ edu/college/corank.htm*. [**Assignment 2**: Find how your college or alma mater fares in the *U.S. News* ranking.] While these rankings are quite controversial and their methodology can be debated at length, I prefer to investigate the extent to which colleges advertise their *U.S. News* rankings. The magazine makes it very easy for colleges to trumpet their success in its publication by providing a "badge" that can be copied from its web site and pasted onto the college's site. I decided to see how many colleges ranked highly by *U.S. News* trumpet this achievement on their web site.

My analysis was conducted on July 3–4, 2000 and was based on the 1999–2000 rankings. *U.S. News* classifies colleges by type as liberal arts colleges or universities and by scope as national or regional. I examined the Web sites of the top ten liberal arts schools, both nationally and in the four regions of the country. For each site, I recorded whether or not I found an official *U.S. News* "badge" displayed at the site and whether I could find any mention of the *U.S. News* ranking. For sites that mentioned the ranking, I also kept track of whether the mention appeared prominently on the college's home page or whether I had to follow links to find it. [**Assignment 3**: Before you look at Table 1, do you expect national or regional schools to mention the ranking more often? **Assignment 4**: What proportion of "top ten" schools do you expect to display a *U.S. News* badge; what proportion do you expect to mention the ranking on their web site? **Assignment 5**: Which region (North, South, Midwest, or West) do you expect to mention the ranking most often? Answers appear below.]

My findings are summarized in Table 1.

Regional schools brag about their *U.S. News* ranking much more often than do national schools. None of the national schools displays the badge, and only 1 of 10 mentions the ranking. On the other hand, 22 of the 40 regional colleges examined (55%) present the badge, and 36 of the 40 (90%) mention the honor. Not much variation among the regions exists, although the West is the only region for which all ten schools mention the ranking.

While on this quest, I found that some colleges advertise their stature in other ranking systems. One that caught my eye was the Templeton Guide's listing of "colleges that encourage character development," available at *www.collegeandcharacter.org/guide/*. Curiously, no college that I have attended or taught at made the top 100 of this list. Another ranking system that captured my attention was Yahoo's ranking of "most wired" colleges, available at *www.zdnet.com/yil/content/college/colleges99.html*.

■ Ranking Places to Live

After choosing a college, another important decision facing the student of statistics is deciding where to live. Again, many ranking systems are available to assist this decision, and many of them are available free on the Web. Perhaps the most famous of these lists is the *Places Rated Almanac*, which is updated every four years and supplies an ordered listing of 333 metropolitan areas in the United States arranged according to "livability." An alternative system for ranking places to live is provided by *Money* magazine. A third option, a free on-line service, is available at www.bestplaces.net. This site offers profiles of 1000 cities on thirty variables in categories including housing, costs of living, crime, education, economy, health, and climate. The site allows users to answer questions that determine their own category weights and thus identifies their "ideal" place to live.

To give you a sense for the information provided on these places, Table 2 lists data on six variables for six cities. The variables are: the purchase cost of a three-bedroom home in a good neighborhood; the violent crime rate per 100,000 residents; the average commuting time in minutes; the high school graduation rate; the number of medical doctors per 10,000 residents; the number of days per year with few clouds; and the average high temperature in July. [**Assignment 6:** Based on these variables, which of the six cities listed appeals most to you?] The cities presented are those of the editorial board and staff of this magazine, in alphabetical order: Alexandria VA, Dallas TX, Des Moines IA, Raleigh NC, San Antonio TX, and San Luis Obispo CA. [**Assignment 7**: Use the data provided in Table 2 to try to identify which city is which. Answers appear at the end.]

For the student who wants to plan far in advance, *Money* magazine provides a "best places to retire" analysis at *www.money.com/money/depts/retirement/bpretire/winners.html*. [**Assignment 8**: Try to identify at least one of their five "winners" for best places to retire. Hint: The states represented by these five cities are Colorado, Florida, Maine, North Carolina, and Oregon. Answers appear at the end.]

Ranking Books

One common element of all students' lives is books. Rankings of books have been around for quite a while, the most famous being the *New York Times* Best-Seller lists. In keeping with the Web theme of this column, though, I set out to find the statistics book with the lowest (i.e., most popular) sales ranking at *amazon.com*.

I thought I had found a winner when I discovered that the classic *How to Lie with Statistics* came in (on July 4, 2000) with a ranking of 7010. I was intrigued to read that two of the places in which this book is particularly popular are Pleasanton, California where it ranks #13 and Northfield, Minnesota where it ranks #5. Upon examining Pleasanton's "top ten" list I found that the Barron's AP Statistics review book ranks #7 in that community, so its residents must be quite statistically literate. Fortunately, I did not abandon my search at this point, for I soon found a statistics book with an even lower sales rank at *amazon.com*. *The Cartoon Guide to Statistics* achieved a rank of 2142. While this does not qualify for *Harry Potter* popularity status, I was surprised to find a statistics book with such a low sales rank. [**Assignment 9**: Try to find a statistics book with a sales rank at amazon.com lower than 2142. Please send me the information if you succeed.]

Table 1.

Badge displayed
Home page mention
Mention

Ranking Presidents

I admit that this one is not especially relevant to students of statistics, but I was intrigued by a ranking of U.S. presidents conducted by C-SPAN in celebration of President's Day 2000. The rankings were compiled through a survey of historians and presidential scholars, and a separate set of rankings was determined through viewer responses. Presidents were rated on ten qualities of presidential leadership, and results can be found at *www.americanpresidents.org/survey/*. [**Assignment 10**: Guess which president had the highest overall rating based on historians' responses and on viewers' responses, and do the same for the lowest overall rating. Answers appear at the end.]

The boxplots (Fig. 1) present the ratings on the ten individual categories for the six most recent presidents: Nixon, Ford, Carter, Reagan, Bush, and Clinton. A category rating is determined by averaging responses from all of the historians. [**Assignment 11**: Before reading on, identify which president goes with which number. Which president from this list do you suspect to have the highest overall rating? Which two do you expect to have the most variability across categories? Who do you think has the outliers on both the low and the high ends? Answers appear in the next paragraph.]

It turns out that Nixon is #5, Ford is #2, Carter is #1, Reagan is #6, Bush is #3, and Clinton is #4. Nixon has outliers on each end, a high rating for international relations and a low rating for moral authority. Reagan has the highest median score among these presidents, while he and Clinton have the most variability. Ford and Bush have the least variability in rating scores among categories of the survey for this group.

While the natural thing is to compare overall, or mean, rating scores among the presidents, I find it interesting to consider the variability in ratings across categories among all 41 presidents. It turns out that the smallest standard deviation was achieved by Rutherford Hayes, followed by Chester Arthur and then Zachary Taylor. The honor of attaining the largest standard deviation goes to Bill Clinton, followed by Ronald Reagan and then Andrew Jackson.

Table 2.

City
Home cost
Violence
Commute
Grad rate
Doctors
Sunny

Ranking Vacation Spots

Having studied colleges and places to live and

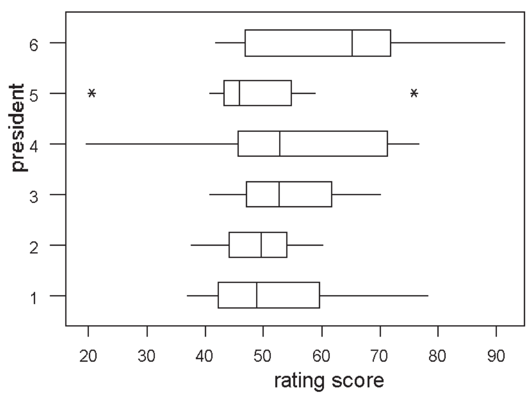


Figure 1.

retirement spots and books and presidents, you're probably ready for some rest and relaxation. What better way to unwind than to head to a beach! As luck would have it, I had the television on in the background as I was writing this column, and the Travel Channel presented a show on the "top ten beaches" in the U.S. In alphabetical order, these beaches are Cape May NJ, Hilton Head SC, Kapalua HI, Panama City FL, Provincetown MA, Santa Barbara CA, South Beach FL, South Padre Island TX, Venice Beach CA, and Waikiki HI. [**Assignment 12**: Try to arrange these in order from best to tenth best as presented by the Travel Channel. Answers appear at the end. **Assignment 13**: Visit *http://travel.discovery.com/ideas/ nature/ beachgd/worldsbest/bestmain.html* to see how many American beaches made the Travel Channel's list of the top ten beaches in the world. **Assignment 14**: Send me a message telling me about your favorite beach, or president, or book, or place to live, or college. **Assignment 15**: Write to ASA headquarters and suggest that this columnist be sent on an all-expenses-paid trip to visit the top ten beaches and to report back on his findings.]

The promised answers are:
1. As you no doubt surmised, this is my fourth *Outlier…s* column, so it should have no difficulty ranking in the top four.
7. 1: San Luis Obispo, 2: Des Moines, 3: Alexandria, 4: San Antonio, 5: Dallas, 6: Raleigh.
8. Fort Collins CO, Bradenton FL, Brunswick ME, Asheville NC, Bend OR.
10. The top overall ranking from historians went to Abraham Lincoln, the lowest to James Buchanan. From viewers, the top and bottom ratings also went to Lincoln and Buchanan, respectively.
12. From best to tenth best: Kapalua, South Beach, Hilton Head, Panama City, Santa Barbara, Provincetown, Waikiki, Cape May, Venice Beach, South Padre Island.

Please send your assignments, and suggestions for future columns, to Allan Rossman at rossman@ dickinson.edu.

Get FIRED up for the 2001
JOINT STATISTICAL MEETINGS in HOT'LANTA

JSM2001
ATLANTA
statistical science for the information age

August 5–9, 2001
Atlanta Hilton & Towers and
Atlanta Marriott Marquis Hotels

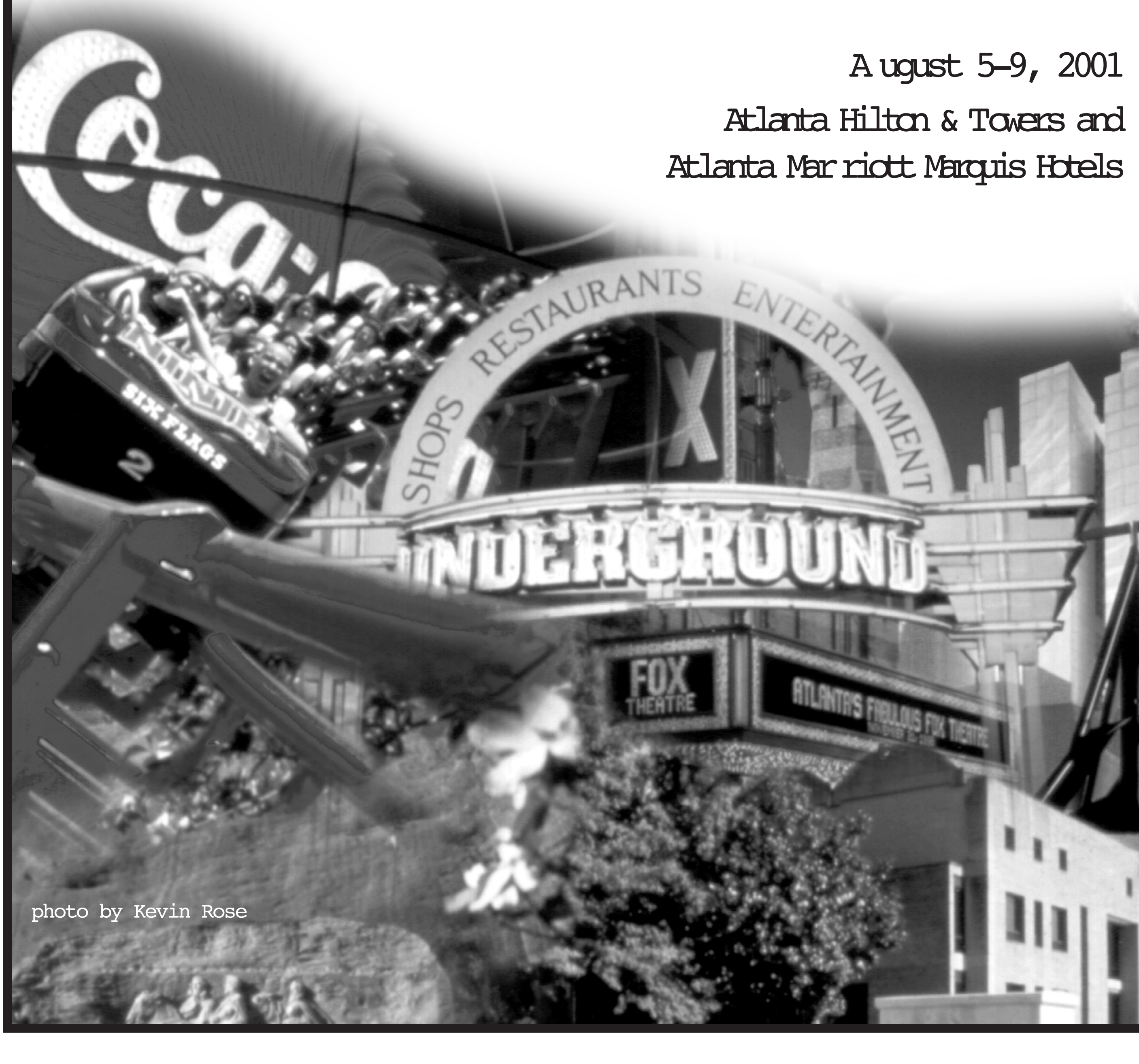


photo by Kevin Rose